

# Psychological Review

EDITED BY

CARROLL C. PRATT  
PRINCETON UNIVERSITY

---

## CONTENTS

- Contributions to Role-Taking Theory: I. Hypnotic Behavior:*  
THEODORE R. SARBIN 255
- Informal Social Communication:* LEON FESTINGER ..... 271
- Dynamic Systems, Psychological Fields, and Hypothetical Constructs:*  
DAVID KRECH 283
- A Quantitative Derivation of Latent Learning:* JAMES DEESE ..... 291
- An Ideal Equation Derived for a Class of Forgetting Curves:*  
IVAN D. LONDON 295
- Simple Qualitative Discrimination Learning:* CLARK L. HULL ..... 303
- A Note on Depth Perception, Size Constancy, and Related Topics:*  
HAROLD SCHLOSBERG 314

---

PUBLISHED BI-MONTHLY BY THE  
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.  
PRINCE AND LEMON STS., LANCASTER, PA.  
AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.  
\$5.50 volume \$1.00 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-3), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

# EARLY PUBLICATION

## IN

# APA JOURNALS

---

The policy of accepting articles for immediate publication (providing the editor accepts the article and the author is willing to pay the entire cost of increasing the next available issue by enough pages to add his article to the normal content) is now standard practice for all APA journals except *Psychological Abstracts* and the *American Psychologist*.

The actual charge made to the author includes three items:

1. A basic charge of so much per page. This is the minimum amount that it costs to add an additional page to the journal. For 1950 these costs are:

	PER PAGE
JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY .....	\$14.00
JOURNAL OF APPLIED PSYCHOLOGY .....	12.00
JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY .....	10.00
JOURNAL OF CONSULTING PSYCHOLOGY .....	11.00
JOURNAL OF EXPERIMENTAL PSYCHOLOGY .....	11.00
PSYCHOLOGICAL BULLETIN .....	16.00
PSYCHOLOGICAL MONOGRAPHS: GENERAL AND APPLIED .....	14.00*
PSYCHOLOGICAL REVIEW .....	12.00

\* Since each Psychological Monograph is printed separately, the author of one handled on an early publication basis can be charged exactly the cost of printing. The figure of \$14.00 is an approximate one; the actual figure will be higher for very short monographs and lower for very long ones. The cost will also vary depending upon the amount of special composition and the illustrations used.

These charges are based upon several factors:

- (a) The greater number of words on a particular journal page, the higher the cost per page. Conversely, the fewer words printed on the page, the lower the cost per page.
  - (b) The more copies which must be printed, the higher the cost.
  - (c) The more expensive the printer, the higher the cost. Compared to the factors listed above, this is not an important difference in the charges made.
2. The full cost of any cuts or other illustrative material, of special composition for tables, and of author's changes in proof.
  3. The full cost of any reprints the author cares to have. (Authors of early-publication articles do not receive any free reprints.)

---

Proceedings of the annual meetings of regional psychological associations are published in the *American Psychologist* at a charge of \$23.00 per page plus the cost of author's alterations and reprints.

# THE PSYCHOLOGICAL REVIEW

## CONTRIBUTIONS TO ROLE-TAKING THEORY:

### I. HYPNOTIC BEHAVIOR <sup>1</sup>

BY THEODORE R. SARBIN

*University of California*

This paper attempts to construct from a social psychological standpoint a workable theory of hypnosis. Briefly stated, it essays to demonstrate that hypnosis is one form of a more general kind of social psychological behavior, namely, role-taking.

That a theory based on social psychological considerations is necessary arises from the obvious social psychological nature of the hypnotic situation. The patent dependency of hypnosis on interpersonal relations calls for a theory which is more continuous with social psychological formulations than with outworn physiological speculations (25) or revived mentalistic entities (46). Moreover, the search for shorter and more efficient psychotherapeutic measures (together with the former widespread use of hypnosis in the treatment of the hysterics) suggests a reconsideration of hypnosis in the treatment of certain behavior disorders. Such treat-

ment will be less abused if it rests on a more substantial theoretical framework than formerly. In addition, the potential value of hypnosis as a tool for social science and medical research demands a careful evaluation of the nature of hypnosis. Thus appropriate allowances will be made for the perturbations in the experimental field introduced by the use of hypnosis as a research instrument.

#### OBSERVATIONS WHICH MUST BE ACCOUNTED FOR

A theory of hypnosis must account for many phenomena subsumed under a single label. These phenomena and the conditions which elicit them may be grouped for our purposes into these four classes: (1) the apparent discontinuity or dissociation of behavior; (2) the apparent automaticity of response; (3) the disjunction between the magnitude of the response and the procedure which instigates the response; and (4) individual differences in responsiveness to hypnotic induction procedures. These four types of observations are briefly elaborated below.

*Apparent discontinuity.* In hypnosis the subject appears to be in a state which is discontinuous from events prior to the initiation of the hypnotic induc-

<sup>1</sup> A preliminary form of this paper was read at the 1946 meetings of the Western Psychological Association. Most of the experimental and clinical work reported in this paper was begun during the author's tenure as a post-doctoral fellow of the Social Science Research Council, 1941-43. The author expresses his gratitude to his colleague, Dr. Harrison G. Gough, and to Dr. R. W. White of Harvard University for critically reading the manuscript.

tion procedure. From introspective accounts and from observers' protocols it seems that stimuli are perceived by a markedly altered organism and that the responses are quantitatively and qualitatively different from those in the pre- and post-hypnotic periods. Some of the more dramatic items of conduct which lead to the acceptance of the inference that the subject's behavior is discontinuous (dissociated) are: anesthesia, amnesia, post-hypnotic compulsive behavior, hypermnesia and various somatic effects such as the inhibition of gastric contractions. To those who are content only with a superficial examination of hypnotic phenomena it appears that hypnotic subjects can perform acts which violate the limits of everyday behavior. When the data are inspected more closely, however, we find that the changes in behavior which do occur involve chiefly the skeletal musculature—i.e., voluntary responses. Responses which are involuntary, such as PGR, blood pressure shifts, and pupillary reflexes are less amenable to verbal instructions, and the limits are extended not too far from the limits of waking behavior (43). Later we shall show that those responses involving the skeletal musculature require no further explanation than that the subject is taking the role of the hypnotic subject as understood by him as a result of his previous interactions with similar social psychological situations. The extension of the limits of behavior involving the autonomic functions is understood in terms of the conception of the organism as-a-whole—a conception which is now generally accepted in sophisticated psychological theory.

*Apparent automaticity.* Most of the early theorists were thrown off the trail of a really workable theory of hypnosis by the manner in which acts are carried out under hypnotic stimulation. The word "trance" has been used to express

this meaning. In most instances the subject appears to act like an automaton. There is an apparent absence of volitional activity. The experimenter throws out commands which seem to be accepted by the subject without critical consideration. He is often slow, stuporous, and seems to be exerting a great deal of effort to perform simple acts. Retrospective accounts reveal a distinction between obedience as found in everyday behavior and the automatic acceptance of commands without the subjective experience of intent. In addition to accounting for this apparent automaticity, a workable hypnotic theory must account for many acts which are added spontaneously by the subject without the benefit of instruction from the experimenter. Unlike physiologically-oriented theories, the role-taking theory considers these observations under the concepts of role-enactment and role perception.

*The disjunction between the magnitude of the response and the procedure which instigates the response.* This aspect of hypnosis is probably responsible for the popular association of hypnosis with magic. The experimenter (or therapist) merely talks to the subject. How, then, can such marked changes in behavior occur merely as a result of verbal instructions? The need for explaining this observation would be less urgent if the stimuli were of the same order of magnitude as are found in extreme stress, fatigue, toxicosis, narcosis, or febrile conditions. In a later section we shall point out how verbal instructions may help the subject focus on and enact a role which may have markedly altered somatic components.

*Individual differences in response to hypnotic induction procedures.* The observation which has received the least attention from the theorists and experimenters is (at least to this writer) the most obvious one, viz., individual sub-



jects respond differently to the same hypnotic procedures. As is well known, many subjects cannot be hypnotized at all, some will exhibit mild cataleptic reactions, and still others will exhibit all the classical responses of hypnosis. Furthermore there is a great deal of variation in the manner in which directions are accepted (or rejected) by subjects who are apparently hypnotized to the same degree. As anyone who has taken the role of a hypnotist knows, and as Brenman (7) has concluded from her analysis of various induction procedures, little or no relationship exists between the subject's performance and the specific innovations which are introduced into the hypnotic instructions. Since the induction procedure *per se* cannot account for the differential responsiveness of subjects, this leaves the subject as a person as the more fruitful focus of study.

These four types of observations may be combined into a question, the answer to which will provide us with a more definitive theory of hypnosis: What are the characteristics of those individuals who, in response to hypnotic induction procedures, exhibit conduct which is apparently discontinuous and apparently automatic?

#### SOME CONCURRENT THEORIES

It is unnecessary to take time out to flog the dead horse of dissociation theory. Numerous experiments and sophisticated observations have led to the unmistakable conclusion that the hypnotized subject is not composed of various psychophysiological systems that can be dissociated one from the other. White and Shevach (45) have written a thoroughgoing analysis of the concept of dissociation and have concluded that the natural cleavages in the nervous system postulated by Janet are nonexistent.

A number of writers cling to the con-

ditioned response theory to explain hypnosis. Historically the conditioned response theory stems from this simple explanation: The word is the conditioned stimulus and acts as an efficient stimulus. This is no more than a streamlining of the old ideomotor hypothesis. In 1933 Hull stated it this way: "... the withdrawal of the subject's symbolic activities would naturally leave his muscles relatively susceptible to the symbolic stimulation emanating continuously from the experimenter . . ." (21, p. 397). From such a conclusion (which seems naively to regard the subject as a spinal animal) Welch has recently presented an hypothesis and an experiment which purport to give credence to the conditioning theory (9, 42). Taking as his point of departure the most commonly used induction procedure, Welch says:

"If the subject analyzed himself in some naïve fashion, he might say, 'When the hypnotist said I felt A, I felt A; when he said I felt B, I felt B; and now he says I feel X, I feel X.' At this point the generalization has extended to the point that whatever the hypnotist says the subject feels, he, within limits, actually feels" (42, p. 361).

On the basis of his hypothesis that hypnosis is a kind of generalized conditioning, Welch and his co-workers performed a learning experiment (in which, incidentally, none of the subjects was hypnotized) based on this experimental analogue. "... a word flashed on a screen was used as analogous to the spoken word of the hypnotist, and followed by the phenomenon for which the word was a symbol. Thus the word 'music' was followed by the playing of music. After a certain number of trials the word 'electric shock' was flashed on the screen and was not re-inforced." His findings were summarized thus: "... in a group of 15 subjects, 11, or

73 per cent gave a (PGR) response greater than to any other stimuli."

That Welch has demonstrated a type of abstract conditioning is not to be denied. But he has not shown that this type of conditioning is the important feature of hypnosis. In the first place, many subjects can be hypnotized without using the analogous procedure. If a subject comes into a hypnotic experiment with certain self-perceptions and role-taking skills, it is possible for him to become hypnotized without the usual monotonous delivery and so-called reinforcement. In an unpublished study (36) the present author has shown that some subjects can be hypnotized with these instructions: "Make yourself comfortable in this easy chair. I'll step out of the room for a few minutes so you can relax. When I come back I will count to ten, you will close your eyes and go into a hypnotic sleep." Even if we could accept the analogy between the Welch experiment and hypnosis, there is no answer to the question: Why did the other 27 per cent not condition? If Welch could show that a correlation existed between "abstract conditionability" and hypnotizability, we should still have to fit this correlation into a more comprehensive framework based on an understanding of the antecedents of these individual differences.

Eysenck and Furneaux (12, 13, 17) have also reported some studies which are related to the ideomotor principle. Using a factorial approach, they isolated three factors from a series of psychomotor and other tests. The first, primary suggestibility, is highly correlated with hypnotizability and is best measured by the postural sway test. The second factor, secondary suggestibility, is unrelated to hypnotizability. The third factor, unrelated to the previous two, also predicts susceptibility to hypnosis, and is measured by a test of heat illusion. They conclude that sus-

ceptibility to hypnosis is an innate characteristic (presumably on the grounds that psychomotor traits are inborn). This writer would declare this conclusion a *non sequitur*. That hypnotizability and certain traits are shown to be related is an acceptable conclusion, but to posit that this relationship is based on inherited factors is not continuous with the data. Below we try to fit these data into our conceptual framework.

Perhaps the most widely accepted hypothesis at the present time is a conative one which places the phenomena of hypnosis at a high integrative level. A number of writers have contributed evidence to support such a theory, notably Dorcus (10), Lundholm (28), Rosenow (32), Pattie (31), White (43), and Sarbin (37). The most systematic presentation of this hypothesis has been offered by White. He defines hypnosis as "meaningful, goal-directed striving, its most general goal being to behave like a hypnotized person as this is continuously defined by the operator and understood by the subject." This approach purports to look upon the hypnotic subject as a functionally intact human organism who is very much in contact with stimulus objects and events, trying to conduct himself in certain meaningful ways rather than in the manner of a spinal animal.

White's theory deals with three of the previously identified four sets of observations. It looks first upon the apparent automaticity as a form of striving: the subject tries to behave in an organized manner, following instructions as he understands them. The apparent discontinuity is treated in terms of measurable extensions of the boundaries of volitional control. How the goal-directed striving makes possible this extension of the limits is subject to speculation in terms of "disinhibition of the

higher centers." The importance of the procedure for inducing hypnosis is analyzed in terms of relaxation, reduction of sensory input, drowsiness, and a contracted frame of reference. This procedure produces an altered state of the organism which makes possible the success achieved by the striving. The theory fails to provide an explanation for differential susceptibility beyond that due to motivational factors, such as need for submissiveness and deference.

This analysis places the striving in a context beginning with the experiment itself. It fails to recognize explicitly that the subject comes into the hypnotic situation with certain pre-conceptions about the experiment, the experimenter, and even about such items as the place in which the experiment is being conducted. It does not make clear that the subject also comes into the hypnotic setting with certain self-perceptions, and that these self-perceptions will operate toward the subject's being successful or not in his striving to behave "in ways defined by the operator." White's analysis would be more tenable if there were no individual differences in responding to the operator's instructions. Relaxation, drowsiness, and reduction of sensory input—time-consuming processes—obviously would not be involved with those subjects who responded immediately to the command: "Go into a hypnotic sleep."<sup>2</sup> The observable dif-

ferences in individuals, not only in the depth of hypnosis, but also in the kind and quality of spontaneous additions to the operator's directions, suggest that we look into the reactional biography of the subject and into the evolution of the stimulus setting for clues as to the nature of hypnosis.

#### THE ROLE-TAKING HYPOTHESIS

To fill the gap in White's goal-striving theory, another hypothesis is herewith introduced. Hypnosis is a form of a more general kind of social psychological behavior known as role-taking. In the hypnotic experiment the subject strives to take the role of the hypnotized person; the success of his striving is a function of favorable motivation, role-perception, and role-taking aptitude. This orientation breaks completely with the tradition of looking on hypnosis as some strange phenomenon for which it is necessary to invent psychophysiological constructions. Rather it is placed in continuity with other social psychological conceptions.<sup>3</sup>

deep vertical activation, reaching to the affective and autonomic levels, of those processes which are suggested. In contrast to this would be the relatively horizontal activation of everyday life where different processes tend to act together or check each other.

"This (monoidism) appears to me to be the pre-dynamic form of what now looks like the best hypothesis for the nature of the hypnotic state. For present purposes some such term as monomotivation would be more suitable. This view of the matter makes possible a fruitful comparison between hypnosis and other states, such as great fear or excitement, in which volition is transcended. All such states are monomotivational but in the sense that one extremely powerful motive or one strong preoccupation momentarily towers over all other processes. Hypnosis achieves the same relative effect at low dynamic intensities, quieting the competitors rather than heightening the chief process."

<sup>3</sup> The concept of role-taking has been described in a previous paper (34). In brief, role-taking may be summarized as follows: (1) Role-enactment depends upon prior ex-

<sup>2</sup> In a personal communication, R. W. White has extended his theory as follows: "It would have been better, I think, to develop at more length the idea of a contracted frame of reference, or, as I would now prefer to put it, a contracted frame of activation. What has to be explained is how the hypnotic suggestions achieve their peculiar success, and I think the explanation should include two things: first, the presence of a single ruling motivation, and second, the exclusion (by quieting) of all promptings and even of the sensory avenues to such promptings that might set up competing processes. In this contracted field of activation there may be conceived to take place a

To adopt a frame of reference that departs from dependence on traditional formulations, and to provide a logical link between the observations and theory, we point to another area of conduct which is apparently automatic, apparently discontinuous, elicited by relatively simple verbal instructions, and characterized by individual differences in performance: to wit, the drama. Introspective accounts and observers' reports of stage actors taking roles reveal a kind of behavior which may be characterized in much the same way as hypnosis. The apparent discontinuity, for example, has been established as an important factor in dramatic role-taking. The actor's stage behavior appears to be dissociated or discontinuous from his "normal personality." In Archer's classical study of acting (2) some actors report losing themselves completely in certain roles so that they are relatively unaware of the audience or of other physical or social objects. The role may even carry over to off-

perience, either symbolic or overt, in order to build up a perception of a given role. (2) Role-taking is organismic, that is to say, it embraces the entire organism, not merely the voluntary reaction-systems. (3) Role-taking occurs with various degrees of participation of the self in the role (this may also be described as levels of consciousness). (4) The perception and enactment of roles is variable inter-individually, intra-individually, and culturally—both qualitatively in terms of the role-behaviors that go to make up any given role, and quantitatively in terms of the number of roles available to an individual or group. (5) Role-taking is a complex form of conduct and can be condensed into significant symbols. (6) Role-taking can be understood as coordinate with the self; a self-concept, phenomenal self, self-dynamism, or ego must be postulated in order to understand role-behavior. In fact, any social psychological behavior. To these may be added another item, (7) statuses or positions, which are established in various ways and which define what roles are appropriate and expected. (See also Cameron [8], especially Chapter IV, and a forthcoming book by the writer, *The Psychology of Role-Taking*.)

stage statuses. The introspective accounts of actors taking roles are often undifferentiated from the accounts of hypnotic subjects (36).

Allen cites Oesterreich who collected a number of observations on this point. One such observation is reproduced here: "Martersteig compares the personality of the theatrical character to a self suggested to the actor by hypnosis, and states that the waking remainder of the actor's consciousness (*Bewusstseinsrest*) can observe the actions of the hypnotic self, as though it were another person, at one time feeling anxiety with regard to them, at another time allowing them to have full play" (1, p. 123).

It appears that the stage director stands in the same relationship to the actor as the hypnotist does to the subject. The statuses or positions are defined beforehand, the specific role-behaviors are dictated by the attempts of each participant to validate his status (27). In short, the participants inter-behave with each other in ways that are appropriate to each position—provided, of course, that such interbehavior can be incorporated by each participant in his self-concept. Because acting has not been burdened with the incubus of dissociation or ideomotor theory, we are not amazed at the frequent marked changes in skeletal and visceral behavior which occur merely because the director tells the actor what to do. The analyst of dramatic acting does not seem to be concerned with such pseudo-problems as the search for a one-to-one constancy relationship between the magnitude of the stimulus (the director's verbal instructions) and the magnitude of the response (the complicated verbal, motor, and visceral reactions of the actor).

From this preliminary description we submit that the role-taking of the stage actor and the role-taking of the hyp-



notic subject embody the same characteristics: (a) Favorable motivation—the actor's self-concept and his perception of the part to which he is assigned must be congruent; if it is not, then his performance is unconvincing or he pays a terrific psychological price. (b) Role-perception—the actor must first perceive the role he is to play—this is achieved partly by the actor's own experiences with similar stage or real-life roles, partly by the director's definition of the role. (c) Role-taking aptitude—needless to say, some actors can take a role more completely than others. Compare, for example, the performance of Barrymore as Hamlet with the efforts of a high school senior.

---

states of ecstasy; mystical experiences;  
role and self undifferentiated

---

hysterias

---

hypnosis

---

"heated" acting

---

technical acting;  
role and self are differentiated

Young (46) has criticized such conceptions of hypnosis by saying that the subject is playing a game with himself and with the experimenter. This criticism is invalid because it does not consider an important dimension. In the two types of role-playing there is a quantitative difference along a continuum which we may characterize as the "conscious-unconscious" dimension. We may ask, how conscious is the actor of his surroundings, of stimulus-objects, and of himself as compared with the hypnotized subject? Or, to put it in terms more continuous with the present study, what is the relative degree of participation of the self in the role (or in Mead's terms, of the "I" in the "me")? Some actors and some hyp-

notic subjects become so involved in the role that perception becomes over-focalized and many self-other observations are by-passed. From those studies of acting which have come to this writer's attention, it would seem that there is a great deal of overlap with hypnotic role-taking in this dimension, but there would be, on the average, less participation of the self in the role of actors as compared with hypnotic subjects. Below is a schematization of this dimension of role-taking, in which acting is placed at a relatively high level of differentiation of self from role. The overlapping in the drawing is intentional. Not only is the relationship of acting to hypnosis shown but these forms of

role-taking are placed in a larger setting the better to illustrate what is meant by this dimension.<sup>3a</sup>

In the last few paragraphs we have tried to orient the reader away from the necessity of physiologizing about hypnosis by showing the similarity of hypnosis and acting. Thus we can conceive of hypnosis as being continuous with other social psychological events. At this time we submit certain observations to lend support to the central hypothesis, *viz.*, hypnotic role-taking is dependent on at least three factors—

<sup>3a</sup> This discussion of the role-taking process is given more detailed treatment in a forthcoming article: Sarbin, Theodore R. and Farberow, Norman L. "Contributions to Role-Taking Theory: II. A Clinical Study of Self and Role."

favorable motivation, role-perception, and role-taking aptitude.

**Favorable motivation.** The most complete paper on this topic has been contributed by White (44). He reviews the studies which have attempted to demonstrate the relationship between hypnotizability and motivational variables. The obtained correlations have for the most part not been significantly different from zero. In his own study White finds a small but positive correlation between hypnotizability and the need for deference (.42), and also a small but negative correlation with the need for autonomy (-.42). "... there is a great deal of individual variation in the tendencies which are awakened, so that manifest needs like *passivity, exhibitionism, sex, or aggression* may sometimes occupy the foreground. . . . There is [also] reason to believe that three latent infantile needs sometimes function as motivating forces favorable to hypnosis: the need for *love*, . . . the tendency for *passive compliance*, . . . and the wish to participate in *omnipotence*. . . ." He concludes with this significant statement. "It is doubtful whether the analysis of motivational factors can be pushed further except by the intensive study of the subjects as individuals" (44, p. 161).

In terms which are more continuous with those of contemporary social psychology, White's conclusion may be restated as follows: If the subject's perception of the self (self-concept) and his perception of the role (here, the role of the hypnotized subject) are not disjunctive or incongruent, then he may be said to be favorably motivated.

One example is herewith presented to

facilitate understanding of this formulation. The author gave a lecture and demonstration of hypnosis to a group of undergraduates. The class instructor had previously pointed out (to the author) several students whom he thought would make good subjects. One of these was a young woman of 21 whom he characterized as being dominated by the need for exhibitionism. She had volunteered, along with several others, to be a subject. She responded to the usual induction procedures and served as the main subject to demonstrate the usual signs of hypnosis, catalepsy, rigidity, hallucinations, post-hypnotic compulsive behavior, amnesia, age-regression (to a period when she could only understand and speak another language), etc. At the end of the meeting those subjects who had passed the usual hypnotic tests were asked if they would participate in an experiment in the author's laboratory. She volunteered along with the others. An appointment was made for a week later. She came with some friends at the appointed hour. But instead of being the easily-hypnotized subject of the week before, she was extremely resistant and showed external signs of anxiety and conflict. After about 30 minutes the experiment was terminated. In an interview which followed, the subject said, "I could not understand why, but every time you said my eyes were getting heavier, I would try harder to keep them open. When you said I would cooperate, I seemed to say to myself, 'I mustn't do this.'" Further questioning revealed that when she had discussed the demonstration with her parents, her father had expressed vehement disapproval of her submitting herself to such indignities, and had instructed her not to participate again. At the time, she thought she gave his instructions little attention, but as the time drew near for keeping the appointment, she became

\* The psychoanalytic theories of hypnosis have contributed little to a systematic understanding of hypnosis *except* in the area of motivation. The transference phenomena (14, 38) can be readily translated into the language of social psychology.

more and more anxious. "You know, I always try to please my father."

In this instance we can say that for the first experiment the subject was favorably motivated. Her self-concept (dominated by the need for exhibitionism, if the instructor's appraisal was correct) and the perception of the role of the hypnotized subject were not disjunctive. In the second experiment the self-concept carried another characteristic—of greater valence than the need for exhibition—the maintenance of her father's approval. The role of the hypnotized subject was incongruent with her self-perception, which perception had been modified by interaction with her father. Although she had demonstrated before that she could perceive the role of the hypnotic subject, and could enact it with great fidelity, she could not focus on the role because of her changed self-perception.

In clinical experience this writer has found that as a patient achieves a set of self-perceptions which makes dependency ego-alien, resistance to hypnosis as a therapeutic aid increases. One patient, near the termination of therapy, was faced with blocking involving her school work. This same symptom had cleared up earlier after a few hypnotic sessions. When it was suggested that hypnosis be used as an auxiliary therapeutic technique, she was resistant to the idea. She said: "I know it worked before, but I would rather work this through on a more mature basis." Janet (24) long ago made the same observation, but related it to different concepts.

**Role-perception.** This concept was first introduced by G. H. Mead (29) and later by Moreno (30) in his studies of the psychodrama. In order to enact a dramatic or psychodramatic role, it is necessary for the subject to have a perception of the role. (The words "image" and "preconception" are used by other writers to express the same

idea [22].) Through various media of communication, such as parental instruction, motion pictures, novels, comic strips, radio stories, rumors and folktales, role perceptions are built up.<sup>5</sup> The role of the father, the role of the teacher, the role of the policeman, etc., are built up from interaction with others in the social environment. When the subject enters the hypnotic situation, then, he comes not only with various self-perceptions, but also with various role-perceptions, among them the role of the hypnotic subject. The announcement of the experiment and the directions of the operator serve as stimuli which elicit the perception of the role. The validity of this conception is suggested by at least three kinds of observations: (1) trance states of certain primitive and religious groups, (2) the role-playing of young children, and (3) clinical and experimental studies.

**Trance states.** In many cultures trance states mark a *rite de passage*. As an illustration we cite one of Benedict's studies. She has described how, among the Plains Indians, an individual will experience many of the phenomena, including hallucinations, which are usually subsumed under the term hypnosis. The content of the hallucinations is relatively constant within groups but highly variable between groups. The role of the tranced subject is perceived from interaction with his own group. "The tranced individual may come back with communications from the dead describing the minutiae of life in the hereafter, or he may visit the world of the unborn, . . . or get information about coming events. Even in trance the individual holds strictly to the rules and expectations of his culture, and his experience is as locally patterned as a

<sup>5</sup> In a paper now in preparation the author analyzes in greater detail how the established principles of perception may be applied to role-perception.

marriage rite or an economic exchange" (6, p. 77). In brief, the perception of the trance role is built up in social interaction.

*Role-playing of young children.* Space prevents the identification of the numerous studies which have been reported dealing with the fantasy-roles observed in young children. One can condense the findings for the purposes of this paper into this general statement: The roles which emerge in the fantasy and play activities of young children are dependent upon their being able to perceive other-roles (4, 5, 8, 15). Some of the studies of imaginary companions are especially illuminating (18).

Hartley *et al.* have recently reported a pioneering study in an attempt to understand how children perceive ethnic group roles and parental roles. As might be expected, children begin to have role-perceptions at an early age and there are levels of complexity in their formulations of role-perception (20).

*Clinical and experimental studies.* Dorcus *et al.* (10) have reported a study which shows clearly that college students—who make up most of the experimental population—are not naïve subjects as far as hypnosis is concerned. For example, of 669 students questioned, 79 per cent answered yes to the question: Is hypnosis possible? To the question, Could you be hypnotized?, 36 per cent said yes, and 15 per cent answered in the affirmative in regard to the possibility of hypnotic amnesia. These data may be interpreted to signify that most college students (the usual experimental population) have a perception of the role of the hypnotic subject. Not all who have such a role-perception, however, can enact the role. The proportion of college students who are successfully hypnotized is much less

than would be expected from the Dorcus *et al.* data.

In an unpublished study (36) the author asked a sophomore class to write descriptions of what takes place in hypnosis. This assignment was made a week before the lecture and demonstration of hypnosis. Volunteers from this class were subject to the induction procedure described by Friedlander and Sarbin (16). The spontaneous acts, introduced by the subjects without instructions from the experimenter, were noted. Of the 12 subjects who volunteered, six subjects were classified as "good" subjects. The spontaneous additions of four of these subjects could have been predicted from their descriptions of the week before. For example, one subject spontaneously awakened from the trance each time she was given a task which called for opening her eyes. Upon a later perusal of her paper, we read "A person's eyes must be closed in order to be in a hypnotic trance." Another subject was non-hypnotizable on the first attempt. On the second trial he performed all the classical tests. His role-description contained the statement: "It takes time to learn to be hypnotized. Most people can't be hypnotized the first time." A third subject performed all the tests satisfactorily, except where she was asked to rise from her chair and write on the blackboard. She was resistant to all suggestions when on her feet. Her paper contained this statement: "The subject has to be reclining or sitting." The fourth subject was extremely stuporous, slow-moving, and unable to perform any of the tests. He required a vigorous shaking in order to wake him from the trance. His paper contained the sentence: "Hypnosis is like a deep sleep, the hypnotizer talks in a low voice and you go into a deep sleep." Of the remaining six subjects, all had a correct perception of the role. Their failure to



enact it could be attributed either to unfavorable motivation or to a lack of role-taking aptitude (*v. infra*). These observations lend support to the notion that variations in role-perception influence role-enactment.

In a clinical study of 10 adult patients in a hospital ward, a standard hypnotic procedure was used except that the operator avoided any mention of the word hypnosis or trance. The words relaxation and restful state were substituted. By any of the usual criteria none of these patients was hypnotized. Five of them fell asleep, however. Later the same subjects were told that hypnosis was to be attempted. They were told about the phenomena of hypnosis, the manner in which it is induced, and the possible therapeutic outcomes. The same induction procedure was used as before but the words hypnosis and hypnotic trance were reinstated. Three of the ten subjects responded to the usual hypnotic tests. Thus certain conditions leading to the perception of the role were prerequisite for enacting the role of the hypnotized subject.

*Role-taking aptitude.* Since motivational factors are necessary but not sufficient to account for the phenomena of hypnosis, and since role-perception does not automatically lead to role-enactment, a role-taking aptitude is postulated. However, since it is impossible to separate the motivational from the aptitudinal factors in studying hypnosis, White has suggested an experimental design (44). To a certain extent this design controls the factor of motivation and allows for an approximate isolation of the hypnotic aptitude. White recommends that all completely unhypnotizable subjects be eliminated for the reason that subjects with unfavorable motivations will thereby be discarded. The remaining subjects may be placed in two groups—sommambu-

lists, showing marked amnesia, hallucinations and anesthesia, and light trance subjects who show eyelid and limb catalepsy. "It can be postulated that the first group possesses the hypnotic aptitude to a marked degree, the second to a moderate degree. There should accordingly be significant differences between their average scores on tests which measure the hypnotic aptitude." This design was adopted in a study conducted at the University of Chicago by the author on an original sample of 70 undergraduate volunteers. All were given the Minnesota Multiphasic Personality Inventory. All were subject to the same induction procedures. Of the 70, 36 were discarded as non-hypnotizable subjects. All verbalized a role-perception (variations in role-perception were not considered). Of the remainder, 16 fell into the category of somnambulistic subjects, and 18 in the category of light trance subjects. Of the various scales on the test, the Hy (hysteria) scale differentiated the two groups. Using a T-score of 55 as a cutting point, the following four-fold table depicts the results.

	Sommambulists	Light trance
55 and above	12	4
Below 55	4	14

The chi-square value is significant to .01. (The mean T-score of the somnambulists was 60, of the light trance subjects, 51.) Thus a scale which differentiates hysterical patients also differentiates hypnotic subjects. This finding recalls that part of Charcot's theory which regards hypnosis as an artificially induced hysteria. However, none of the subjects was known to be a hysterical patient. We are led to the same conclusions made by clinicians for many years—the good hypnotic subject and the hysterical patient have something in common. We would suggest the role-taking aptitude.

Auxiliary support is given to this conclusion in a study reported by Lewis and Sarbin (25). Here hypnotic subjects were told to imagine eating a meal at a time when they were having gastric hunger contractions. We found a high correlation between the depth of hypnosis (Friedlander-Sarbin scale) and the ability to inhibit hunger contractions. Those who could take the role of the eater—to use an expression of Moreno's—who could imagine themselves ingesting food, initiated a set of internal responses which resulted in the cessation of the gastric contractions. Subjects who could not be hypnotized, who could not take the role either of the hypnotic subject, or of the eater in imagination, showed no cessation of gastric contractions. That role-taking is organismic is demonstrated here.

When we say that the role-taking aptitude is organismic we refer back to our "observations which must be accounted for." We repeated the question raised by the laity and by other theorists: How can such marked changes in behavior result from such apparently innocuous stimuli? <sup>7</sup> It is probably not far from the truth to say with Goldstein (19) that any act involves the entire organism. When an individual places

himself in the hypnotic situation—when he takes the role of the hypnotic subject—he does so organismically. When the subject acts *as if* he is ingesting food, his actions are total. The variation in his bodily responses, of course, will vary with the completeness and intensity of the role-taking.

A further comment is required about the organismic basis of the role-taking aptitude, especially as seen in acts which transcend normal limits. In the case of actors taking a stage role there are some who will enact the role without a preliminary warming-up process, while others require "preparation." In this warming-up or preparatory process the director helps the actor perceive some of the necessary attributes of the role. This might be considered a kind of covert practice in role-taking. In hypnosis the frequent lengthy induction may serve the same purpose, especially where the subject requires time to shift to the type of attentional behavior which is a component of the hypnotic role. Relaxation, diffuseness, and uncritical passivity as components of the role may be perceived by the subject as a result of the experimenter's instructions. When the subject aptly takes the hypnotic role (whether immediately, or after warming up *via* the induction procedures) a shift occurs from a sharp, alert, objective and critical attitude to a relatively relaxed, diffuse, and uncritical one. Because the alert orientation is highly valued and supported in our society some coaching or "preparation" is required for certain subjects. They must shift their focus to a relaxed, diffuse orientation which (as in the case of mystical states, for example) allows for more active motor-involvement and more intense affectivity. The variations in intensity or completeness with which one takes a role, and the concurrent motor and autonomic effects, are probably related to the subject's ability to

<sup>7</sup> A philosophical digression is in order here. Scientists, no less than laymen, are influenced and limited by their historical and cultural horizons. Growing up in an intellectual environment in which a dichotomy is made between mind and body, between mental events and physical events, scientists are "amazed" when they observe events which are not congruent with the dichotomy. When a scientist's *eidōs* is freed from the necessity of fractionating behavior into the dichotomy dictated by 17th century dualism, then he can regard human behavior as organismic. Why should social psychological events not serve as conditions for altering predominantly biological activities? No one is amazed when respiratory changes are observed in attention experiments, or BMR's of westerners become more like those of orientals when living under specified oriental conditions, etc.

utilize *as-if* formulations. It is to this notion that we now turn.

#### THE *As-If* FORMULATION

Upon what does the role-taking aptitude depend? In a prior paragraph we noted the apparent relationship between the role-taking of the drama and role-taking in hypnosis. Mr. Arbuthnot, the actor, in taking the part of Hamlet, acts *as if* he is Hamlet and not Mr. Arbuthnot. The hypnotic subject acts *as if* he is an automaton (if automaticity is included in his role-perception). As a preliminary postulate we can say that the role-taking aptitude depends upon the subject's participation in *as-if* behavior. That this has a more general application is seen from a logical analysis of Rosenzweig's "triadic hypothesis" (33). In this statement, hypnotizability as a personality trait, repression as an ego-defense, and impunitiveness as a response to frustration are shown to be related. These may be considered *as-if* structures. We have already noted the *as-if* character of hypnosis. In repression the subject acts *as if* an event threatening to the self had not occurred. In the impunitiveness response to frustration, similarly, the subject acts *as if* the frustrating event were no longer frustrating. The *as-if* formulation may be seen not only in the drama, in hypnosis, but in fantasy, play, and, in fact, all imaginative behavior. Imaginative behavior is *as-if* behavior (40). Some data have been put forward by Jacobson (23), Schultz (39), Arnold (3), and others which may be put to use in formulating our theory. From the proposition that all imaginative behavior is *as-if* behavior, we may state that role-taking aptitude depends upon imagination. The following statements give at least initial validity to this proposition.

In a series of carefully controlled studies Jacobson (23) was able to dem-

onstrate the influence of the subject's imagining certain events upon bodily functions. For example, in a condition of relaxation, a subject was told to imagine elevating his arm. The electrical recording showed activity in the muscles which were involved. Schultz (39) reports many instances of the influence of imagination on various muscular and vascular characteristics. Varondenck (41) tells how imaginary processes (implicit) can spill over into overt muscular movements during the act of imagining. Common experience verifies the same notion. In imagining a former embarrassing situation we can feel our ears reddening and our faces flushing; in imagining a former painful experience we may involuntarily withdraw from the direction of the imagined stimulus, or in imagining something extremely unpleasant or disgusting we may experience nausea.

Arnold has written the most complete analysis of the relationship between hypnosis and imagination (3). According to her hypothesis, "... in hypnosis the individual is actively striving to imagine what the hypnotist describes, and in so doing gradually narrows down his focus and relinquishes control of his imaginative processes. . . . The individual focuses on a situation and actively selects the sensations which he will perceive; he actively focuses on possible situations in imagining, on symbols in logical thinking; and he refocuses on past experiences in remembering. Such focussing . . . is merely directed more efficiently, more intensely, during hypnosis than in waking life, and determined by the hypnotist instead of by the subject himself" (3, p. 127). This writer would amend the last statement to read: The focussing is determined by the hypnotist only insofar as the subject's self-perceptions and role-perception permit such direction. This amendment would follow from a

careful consideration of the data Arnold presents from her own experiment which reveals the individual character of the subject's own imagining over and above the directions of the experimenter.

Although Arnold's views are more sophisticated than most previous theories, we are left without any anchorage point for understanding differential responsiveness. The numerous experiments cited by Arnold show the influence of imagination on behavior and the kinds of experimental and clinical situations appear to be of the same kind as the hypnotic situation. But what of the answer to the all important social-psychological question: What are the characteristics of those individuals who are *not* able to focus and thus cannot produce changes in overt or covert behavior?

In Arnold's data is concealed a partial answer to this question. She reports an experiment in which the postural sway technique is used. She tested the hypothesis that a suggestion is acted upon only if the subject actively imagines it. The subjects were told to imagine falling forward. The amount of postural sway was recorded. Comparisons were made between the amount of sway and the reported vividness of imagery. Her conclusion was: The more vivid the imaginative process, the more pronounced the overt movements. From this conclusion and from the long-accepted conclusion about the relationship between the postural sway test and hypnotizability a correlation between vividness of imagery and hypnotic depth could be posited. We could then deduce that hypnotic role-taking depended upon imaginative (*as-if*) processes.<sup>8</sup> One might fit the previously

mentioned findings of Eysenck and Furneaux into this formulation. Subjects who score high on postural sway tests and test of heat illusion are able to imagine vividly in these sense modalities. *A fortiori*, the experiment of Sarbin and Madow (37) may be cited in which the depth of hypnosis and the Rorschach W/D ratio were shown to be correlated. The W or Whole response purportedly indicates a more active imagination.

How, then, does the role-taking theory apply to the four sets of observations previously identified as requiring explanation?

The apparent automaticity is apparent only. The subject varies his responses to the hypnotic situation in terms of his perception of the role of the hypnotized subject. If his perception includes automaticity, then he will act like an automaton.

The apparent discontinuity of behavior is also apparent but not real. The subject's behavior is continuous with his pre-experimental behavior—modified only by his enactment of the role of the hypnotic subject. Such "discontinuous" behavior as amnesia, post-hypnotic compulsions, etc., can be understood in terms of the subject's perception of the role, of his facility in *as-if* behavior and of the degree of participation of the self in the role.

The apparent disjunction between the magnitude of the response and the procedure for eliciting the response is a pseudo-problem. The magnitude of the response is not dependent upon the procedure except insofar as it coincides with the role-expectations of the subject. What appears to be a disjunction is a vestigial remnant of an outmoded psychology which sought to find constancy between phenomenal experience and stimulus events. If the subject has an adequate perception of the role, if this perception is not incongruent with

<sup>8</sup> Clinically, the writer has never found an adult with eidetic or vivid imagery who was not a good hypnotic subject. In a personal communication D. W. MacKinnon reports the same observation.



his self-perceptions, and if he has an appropriate amount of the role-taking aptitude, then he will produce all the dramatic phenomena of hypnosis merely because "the operator talks to him." If he does not or cannot perceive the role, if the role is not congruent with his self-perceptions, and/or he does not have a sufficient amount of the role-taking aptitude or skill, then he will not respond to the operator's commands. Thus differential responsiveness is declared to be a function of these three variables.

#### SUMMARY

The known facts about hypnosis were grouped in four classes of observations: (1) apparent automaticity, (2) apparent discontinuity, (3) disjunction between the magnitude of the stimulus and the magnitude of the response, and (4) differential responsiveness. Because of the obvious dependence of the first three factors upon the fourth (differential responsiveness) this question was formulated: What are the characteristics of those individuals who, in response to hypnotic induction procedures, exhibit conduct which is apparently discontinuous and apparently automatic?

We sought to demonstrate that concurrent theories of hypnosis were tradition-bound: trying to explain hypnotic behavior in terms of conditioning, heredity, or vague neurological formulae. In order to establish a logical link between hypnosis and another form of social psychological conduct which is accepted without resorting to traditional formulations, we first indicated the similarity between role-taking in the drama and role-taking in hypnosis. We postulated that success in taking a dramatic role or hypnotic role depended upon favorable motivation, a perception of the role, and role-taking aptitude. The chief difference in the two forms of role-taking was the degree of participation of

the self in the role (levels of consciousness).

The main portion of our presentation attempted to establish the validity of these conceptions. Favorable motivation was re-defined as congruence between the subject's self-concept and the role of the hypnotic subject. Role-perception is derived from the individual's interaction with various media of communication: the manner in which role-perception influences role-enactment is indicated. Finally, a role-taking aptitude is postulated. From our present state of knowledge this aptitude is probably dependent upon or continuous with the ability of the subject to use *as-if* formulations. Various research and clinical findings were introduced to supply a groundwork for the initial validity of the argument.

#### REFERENCES

1. ALLEN, A. H. B. *The self in psychology*. London: Kegan Paul, Trench, Trubner & Co., 1935.
2. ARCHER, W. *Masks or faces?* New York: Longmans, Green & Co., 1889.
3. ARNOLD, M. B. On the mechanism of suggestion and hypnosis. *J. abn. soc. Psychol.*, 1946, **41**, 107-128.
4. AXLINE, V. *Play therapy*. Boston: Houghton-Mifflin Co., 1947.
5. BACH, G. Young children's play fantasies. *Psychol. Monogr.*, 1945, **59**, No. 2.
6. BENEDICT, R. Anthropology and the abnormal. *J. gen. Psychol.*, 1934, **10**, 59-82.
7. BRENNAN, M., & GILL, M. M. *Hypnotherapy*. New York: International Universities Press, 1947.
8. CAMERON, N. *The psychology of behavior disorders*. Boston: Houghton-Mifflin Co., 1947.
9. CORN-BECKER, F., WELCH, L., & FISICHELLI, V. Conditioning factors underlying hypnosis. *J. abn. soc. Psychol.*, 1949, **44**, 212-222.
10. DORCUS, R., BRETNALL, A. K., & CASE, H. W. Control experiments and their relation to theories of hypnosis. *J. gen. Psychol.*, 1941, **24**, 217-221.
11. DYMOND, R. A scale for the measurement of empathic ability. *J. consult. Psychol.*, 1948, **13**, 127-133.

12. EYSENCK, H. J. Suggestibility and hysteria. *J. Neurol. Psychiat.*, 1943, 6, 22-31.
13. —, & FURNEAUX, W. D. Primary and secondary suggestibility. *J. exp. Psychol.*, 1945, 35, 485.
14. FERENCZI, S. *Sex in psychoanalysis*. Boston: Richard G. Badger, 1916.
15. FLUGEL, J. C. *Man, morals, and society*. New York: International Universities Press, 1944.
16. FRIEDLANDER, J. W., & SARBIN, T. R. The depth of hypnosis. *J. abn. soc. Psychol.*, 1938, 33, 453-475.
17. FURNEAUX, W. D. Prediction of susceptibility to hypnosis. *J. Personal.*, 1946, 14, 281-294.
18. GREEN, G. *The day dream*. London: Univ. of London Press, 1923.
19. GOLDSTEIN, K. *The organism*. New York: American Book Co., 1939.
20. HARTLEY, E. L., ROSENBAUM, M., & SCHWARTZ, S. Children's perceptions of ethnic group membership. *J. Psychol.*, 1948, 26, 387-398.
21. HULL, C. L. *Hypnosis and suggestibility*. New York: D. Appleton-Century, 1933.
22. ICHHEISER, G. Misunderstandings in human relations—a study in false social perception. *Amer. J. Sociol.*, 1949, 55, Part 2.
23. JACOBSON, E. *Progressive relaxation*. Chicago: Univ. of Chicago Press (rev. ed.), 1938.
24. JANET, P. *Major symptoms of hysteria*. New York: Macmillan Co., 1907.
25. KUBIE, L. S., & MARGOLIN, S. The process of hypnotism and the nature of the hypnotic state. *Amer. J. Psychiat.*, 1944, 100, 611-622.
26. LEWIS, J. H., & SARBIN, T. R. Studies in psychosomatics: the influence of hypnotic stimulation on gastric hunger contractions. *Psychosom. Med.*, 1943, 5, 125-131.
27. LINTON, R. *The cultural background of personality*. London: Routledge & Kegan Paul, Ltd., 1947.
28. LUNDHOLM, H. An experimental study of functional anesthetics as induced by suggestion in hypnosis. *J. abn. soc. Psychol.*, 1928, 23, 337-355.
29. MEAD, G. H. *Mind, self, and society*. Chicago: Univ. of Chicago Press, 1934.
30. MORENO, J. L. Role tests and role diagrams of children. In *Psychodrama*, New York: Beacon House, 1946, Vol. 1.
31. PATTIE, F. A. The production of blisters by hypnotic suggestions: a review. *J. abn. soc. Psychol.*, 1941, 36, 62-72.
32. ROSENOW, C. Meaningful behavior in hypnosis. *Amer. J. Psychol.*, 1928, 40, 205-235.
33. ROSENZWEIG, S., & SARASON, S. An experimental study of the triadic hypothesis: reaction to frustration, ego-defence, and hypnotizability. *Character & Pers.*, 1942, 11, 1-19.
34. SARBIN, T. R. The concept of role-taking. *Sociometry*, 1943, 6, 273-285.
35. —. Rorschach patterns under hypnosis. *Amer. J. Orthopsychiat.*, 1939, 11, 315-318.
36. —. Studies in role-taking (unpublished).
37. —, & MADOW, L. Predicting the depth of hypnosis by means of the Rorschach test. *Amer. J. Orthopsychiat.*, 1942, 12, 268-270.
38. SCHILDER, P., & KAUDERS, O. Hypnosis. *Nerv. ment. Dis. Monogr. Ser.*, 1927, No. 46.
39. SCHULTZ, J. *Das Autogene Training (Konzentrierte Selbstentspannung)*, Leipzig, 1932.
40. VAHINGER, H. *The philosophy of 'as-if'*. London: K. Paul, Trench, Trubner & Co., Ltd., 1924.
41. VARENDONCK, J. *The psychology of day dreams*. London: George Allen & Unwin, Ltd., 1921. Pp. 367.
42. WELCH, L. A behavioristic explanation of the mechanism of suggestion and hypnosis. *J. abn. soc. Psychol.*, 1947, 42, 359-364.
43. WHITE, R. M. A preface to the theory of hypnotism. *J. abn. soc. Psychol.*, 1941, 36, 477-505.
44. —. An analysis of motivation in hypnosis. *J. gen. Psychol.*, 1941, 24, 145-162.
45. —, & SHEVACH, S. Hypnosis and the concept of dissociation. *J. abn. soc. Psychol.*, 1937, 42, 309-328.
46. YOUNG, P. C. Hypnotic regression—fact or artifact? *J. abn. soc. Psychol.*, 1940, 35, 273-278.

[MS. received January 11, 1950]

## INFORMAL SOCIAL COMMUNICATION

BY LEON FESTINGER

*Research Center for Group Dynamics, University of Michigan*

The importance of strict theory in developing and guiding programs of research is becoming more and more recognized today. Yet there is considerable disagreement about exactly how strict and precise a theoretical formulation must be at various stages in the development of a body of knowledge. Certainly there are many who feel that some "theorizing" is too vague and indefinite to be of much use. It is also argued that such vague and broad "theorizing" may actually hinder the empirical development of an area of knowledge.

On the other hand there are many who express dissatisfaction with instances of very precise theories which do exist here and there, for somehow or other a precise and specific theory seems to them to leave out the "real" psychological problem. These persons seem to be more concerned with those aspects of the problem which the precise theory has not yet touched. From this point of view it is argued that too precise and too strict theorizing may also hinder the empirical development of an area of knowledge.

It is probably correct that if a theory becomes too precise too early it can have tendencies to become sterile. It is also probably correct that if a theory stays too vague and ambiguous for too long it can be harmful in that nothing can be done to disprove or change it. This probably means that theories, when vague, should at least be stated in a form which makes the adding of precision possible as knowledge increases. It also probably means that theory should run ahead, but not too far ahead, of the data so that the trap of pre-

mature precision can be avoided. It certainly means that theories, whether vague or precise, must be in such a form that empirical data can influence them.

This article is a statement of the theoretical formulations which have been developed in the process of conducting a program of empirical and experimental research in informal social communication. It has grown out of our findings thus far and is in turn guiding the future course of the research program.<sup>1</sup> This program of research concerns itself with finding and explaining the facts concerning informal, spontaneous communication among persons and the consequences of the process of communication. It would seem that a better understanding of the dynamics of such communication would in turn lead to a better understanding of various kinds of group functioning. The theories and hypotheses presented below vary considerably in precision, specificity and the degree to which corroborating data exist. Whatever the state of precision, however, the theories are empirically oriented and capable of being tested.

Since we are concerned with the spontaneous process of communication which goes on during the functioning of groups we must first differentiate the variety of types of communication which occur according to the theoretical conditions which give rise to tendencies to communicate. It is plausible to assume that separating the sources or origins of pressures to communicate that may act

<sup>1</sup> This research program consists of a number of coordinated and integrated studies, both in the laboratory and in the field. It is being carried out by the Research Center for Group Dynamics under contract N6onr-23212 NR 151-698 with the Office of Naval Research.

on a member of a group will give us fruitful areas to study. This type of differentiation or classification is, of course, adequate only if it leads to the separation of conceptually clear areas of investigation within which communication can be organized into storable theoretical and empirical laws.

We shall here deal with those few of the many possible sources of pressures to communicate in which we have thus far been able to make theoretical and empirical progress. We shall elaborate on the theory for regarding them as giving rise to pressures to communicate and on specific hypotheses concerning the laws of communication which stem from these sources.

#### I. PRESSURES TOWARD UNIFORMITY IN A GROUP

One major source of forces to communicate is the pressure toward uniformity which may exist within a group. These are pressures which, for one reason or another, act toward making members of a group agree concerning some issue or conform with respect to some behavior pattern. It is stating the obvious, of course, to say that these pressures must be exerted by means of a process of communication among the members of the group. One must also specify the conditions under which such pressures toward uniformity arise, both on a conceptual and an operational level so that in any specific situation it is possible to say whether or not such pressures exist. We shall, in the following discussion, elaborate on two major sources of pressures toward uniformity among people, namely, social reality and group locomotion.

1. *Social reality*: Opinions, attitudes, and beliefs which people hold must have some basis upon which they rest for their validity. Let us as a start abstract from the many kinds of bases for the subjective validity of such opinions,

attitudes, and beliefs one continuum along which they may be said to lie. This continuum we may call a scale of degree of physical reality. At one end of this continuum, namely, complete dependence upon physical reality, we might have an example such as this: A person looking at a surface might think that the surface is fragile or he might think that the surface is unbreakable. He can very easily take a hammer, hit the surface, and quickly be convinced as to whether the opinion he holds is correct or incorrect. After he has broken the surface with a hammer it will probably make little dent upon his opinion if another person should tell him that the surface is unbreakable. It would thus seem that where there is a high degree of dependence upon physical reality for the subjective validity of one's beliefs or opinions the dependence upon other people for the confidence one has in these opinions or beliefs is very low.

At the other end of the continuum where the dependence upon physical reality is low or zero, we might have an example such as this: A person looking at the results of a national election feels that if the loser had won, things would be in some ways much better than they are. Upon what does the subjective validity of this belief depend? It depends to a large degree on whether or not other people share his opinion and feel the same way he does. If there are other people around him who believe the same thing, then his opinion is, to him, valid. If there are not others who believe the same thing, then his opinion is, in the same sense, not valid. Thus where the dependence upon physical reality is low the dependence upon social reality is correspondingly high. An opinion, a belief, an attitude is "correct," "valid," and "proper" to the extent that it is anchored in a group of



people with similar beliefs, opinions, and attitudes.

This statement, however, cannot be generalized completely. It is clearly not necessary for the validity of someone's opinion that everyone else in the world think the way he does. It is only necessary that the members of that group to which he refers this opinion or attitude think the way he does. It is not necessary for a Ku Klux Klanner that some northern liberal agree with him in his attitude toward Negroes, but it is eminently necessary that there be other people who also are Ku Klux Klanners and who do agree with him. The person who does not agree with him is seen as different from him and not an adequate referent for his opinion. The problem of independently defining which groups are and which groups are not appropriate reference groups for a particular individual and for a particular opinion or attitude is a difficult one. It is to some extent inherently circular since an appropriate reference group tends to be a group which does share a person's opinions and attitudes, and people tend to locomote *into* such groups and *out of* groups which do not agree with them.

From the preceding discussion it would seem that if a discrepancy in opinion, attitude, or belief exists among persons who are members of an appropriate reference group, forces to communicate will arise. It also follows that the less "physical reality" there is to validate the opinion or belief, the greater will be the importance of the social referent, the group, and the greater will be the forces to communicate.

2. *Group locomotion*: Pressures toward uniformity among members of a group may arise because such uniformity is desirable or necessary in order for the group to move toward some goal. Under such circumstances there

are a number of things one can say about the magnitude of pressures toward uniformity.

(a) They will be greater to the extent that the members perceive that group movement would be facilitated by uniformity.

(b) The pressures toward uniformity will also be greater, the more dependent the various members are on the group in order to reach their goals. The degree to which other groups are substitutable as a means toward individual or group goals would be one of the determinants of the dependence of the member on the group.

We have elaborated on two sources of pressure toward uniformity among members of groups. The same empirical laws should apply to communications which result from pressures toward uniformity irrespective of the particular reasons for the existence of the pressures. We shall now proceed to enumerate a set of hypotheses concerning communication which results from pressures toward uniformity.

## II. HYPOTHESES ABOUT COMMUNICATION RESULTING FROM PRESSURES TOWARD UNIFORMITY

Communications which arise from pressures toward uniformity in a group may be seen as "instrumental" communications. That is, the communication is not an end in itself but rather is a means by which the communicator hopes to influence the person he addresses in such a way as to reduce the discrepancy that exists between them. Thus we should examine the determinants of: (1) when a member communicates, (2) to whom he communicates and (3) the reactions of the recipient of the communication.

(1) Determinants of the magnitude of pressure to communicate:

Hypothesis 1a: *The pressure on members to communicate to others in the*

*group concerning "item x" increases monotonically with increase in the perceived discrepancy in opinion concerning "item x" among members of the group.*

Remembering that we are considering only communication that results from pressures toward uniformity, it is clear that if there are no discrepancies in opinion, that is, uniformity already exists in the group, there will be no forces to communicate. It would be plausible to expect the force to communicate to increase rapidly from zero as the state of affairs departs from uniformity.

*Hypothesis 1b: The pressure on a member to communicate to others in the group concerning "item x" increases monotonically with increase in the degree of relevance of "item x" to the functioning of the group.*

If "item x" is unimportant to the group in the sense of not being associated with any of the values or activities which are the basis for the existence of the group, or if it is more or less inconsequential for group locomotion, then there should be few or no forces to communicate even when there are perceived discrepancies in opinion. As "item x" becomes more important for the group (more relevant), the forces to communicate when any given magnitude of perceived discrepancy exists, should increase.

Corroborative evidence for this hypothesis is found in an experiment by Schachter (8) where discussion of the same issue was experimentally made relevant for some groups and largely irrelevant for others. It is clear from the data that where the discussion was relevant to the functioning of the group there existed stronger forces to communicate and to influence the other members. Where the issue is a relevant one the members make longer individual contributions to the discussion and there

are many fewer prolonged pauses in the discussion.

*Hypothesis 1c: The pressure on members to communicate to others in the group concerning "item x" increases monotonically with increase in the cohesiveness of the group.*

Cohesiveness of a group is here defined as the resultant of all the forces acting on the members to remain in the group. These forces may depend on the attractiveness or unattractiveness of either the prestige of the group, members in the group, or the activities in which the group engages. If the total attraction toward the group is zero, no forces to communicate should arise; the members may as easily leave the group as stay in it. As the forces to remain in the group increase (given perceived discrepancies in opinion and given a certain relevance of the item to the functioning of the group) the pressures to communicate will increase.

Data from an experiment by Back (1) support this hypothesis. In this experiment groups of high and low cohesiveness were experimentally created using three different sources of attraction to the group, namely, liking the members, prestige attached to belonging, and possibility of getting a reward for performance in the group activity. For each of the three types of attraction to the group the more cohesive groups were rated as proceeding at a more intense rate in the discussion than the corresponding less cohesive groups. In addition, except for the groups where the attraction was the possibility of reward (perhaps due to wanting to finish and get the reward) there was more total amount of attempted exertion of influence in the highly cohesive groups than in the less cohesive groups. In short, highly cohesive groups, having stronger pressures to communicate, discussed the issue at a more rapid pace and attempted to exert more influence.

(2) Determinants of choice of recipient for communications:

Hypothesis 2a: *The force to communicate about "item x" to a PARTICULAR MEMBER of the group will increase as the discrepancy in opinion between that member and the communicator increases.*

We have already stated in Hypothesis 1a that the pressure to communicate in general will increase as the perceived non-uniformity in the group increases. In addition the force to communicate will be strongest toward those whose opinions are most different from one's own and will, of course, be zero towards those in the group who at the time hold the same opinion as the communicator. In other words, people will tend to communicate to those within the group whose opinions are most different from their own.

There is a clear corroboration of this hypothesis from a number of studies. In the previously mentioned experiment by Schachter (8) the distribution of opinions expressed in the group was always as follows: Most of the members' opinions clustered within a narrow range of each other while one member, the deviate, held and maintained an extremely divergent point of view. About five times as many communications were addressed to the holder of the divergent point of view as were addressed to the others.

In an experiment by Festinger and Thibaut (5) the discussion situation was set up so that members' opinions on the issue spread over a considerable range. Invariably 70 to 90 per cent of the communications were addressed to those who held opinions at the extremes of the distribution. The curve of number of communications received falls off very rapidly as the opinion of the recipient moves away from the extreme of the distribution. The hypothesis would seem to be well substantiated.

Hypothesis 2b: *The force to communicate about "item x" to a PARTICULAR PERSON will decrease to the extent that he is perceived as not a member of the group or to the extent that he is not wanted as a member of the group.*

From the previous hypothesis it follows that communications will tend to be addressed mainly toward those with extreme opinions within the group. This does not hold, however, for any arbitrarily defined group. The present hypothesis, in effect, states that such relationships will apply only within *psychological* groups, that is, collections of people that exist as groups psychologically for the members. Communications will tend not to be addressed towards those who are not members of the group.

The study by Schachter (8) and the study by Festinger and Thibaut (5) both substantiate this hypothesis. In Schachter's experiment those group members who do not want the person holding the extremely divergent point of view to remain in the group tend to stop communicating to him towards the end of the discussion. In the experiment by Festinger and Thibaut, when the subjects have the perception that the persons present include different kinds of people with a great variety of interests, there tends to be less communication toward the extremes in the last half of the discussion after the rejection process has had time to develop. In short, communication towards those with different opinions decreases if they are seen as not members of the *psychological* group.

Hypothesis 2c: *The force to communicate "item x" to a particular member will increase the more it is perceived that the communication will change that member's opinion in the desired direction.*

A communication which arises because of the existence of pressures toward uniformity is made in order to

exert a force on the recipient in a particular direction, that is, to push him to change his opinion so that he will agree more closely with the communicator. If a member is perceived as very resistant to changing his opinion, the force to communicate to him decreases. If it seems that a particular member will be changed as the result of a communication so as to increase the discrepancy between him and the communicator, there will exist a force not to communicate to him. Thus under such conditions there will be tendencies *not* to communicate this particular item to that member.

There is some corroboration for this hypothesis. In a face to face verbal discussion where a range of opinion exists, the factors which this hypothesis points to would be particularly important for those members whose opinions were near the middle of the range. A communication which might influence the member at one extreme to come closer to the middle might at the same time influence the member at the other extreme to move farther away from the middle. We might then expect from this hypothesis that those holding opinions in the middle of the existing range would communicate less (because of the conflict) and would address fewer communications to the whole group (attempting to influence only one person at a time).

A number of observations were conducted to check these derivations. Existing groups of clinical psychologists who were engaging in discussions to reconcile their differences in ratings of applicants were observed. Altogether, 147 such discussions were observed in which at least one member's opinion was in the middle of the existing range. While those with extreme opinions made an average of 3.16 units of communication (number of communications weighted by length of the communica-

tion), those with middle opinions made an average of only 2.6 units of communication. While those with extreme opinions addressed 38 per cent of their communications to the whole group, those with middle opinions addressed only 29 per cent of their communications to everyone.

(3) Determinants of change in the recipient of a communication:

Hypothesis 3a: *The amount of change in opinion resulting from receiving a communication will increase as the pressure towards uniformity in the group increases.*

There are two separate factors which contribute to the effect stated in the hypothesis. The greater the pressure towards uniformity, the greater will be the amount of influence exerted by the communications and, consequently, the greater the magnitude of change that may be expected. But the existence of pressures toward uniformity will not only show itself in increased attempts to change the opinions of others. Pressures toward uniformity will also produce greater readiness to change in the members of the group. In other words, uniformity may be achieved by changing the opinions of others and/or by changing one's own opinions. Thus we may expect that with increasing pressure towards uniformity there will be less resistance to change on the part of the members. Both of these factors will contribute to produce greater change in opinion when the pressure toward uniformity is greater.

There is evidence corroborating this hypothesis from the experiment by Festinger and Thibaut (5). In this experiment three degrees of pressure towards uniformity were experimentally induced in different groups. Irrespective of which of two problems were discussed by the group and irrespective of whether they perceived the group to be homogeneously or heterogeneously com-



posed, the results consistently show that high pressure groups change most, medium pressure groups change next most, and low pressure groups change least in the direction of uniformity. While the two factors which contribute to this effect cannot be separated in the data, their joint effect is clear and unmistakable.

Hypothesis 3b: *The amount of change in opinion resulting from receiving a communication will increase as the strength of the resultant force to remain in the group increases for the recipient.*

To the extent that a member wishes to remain in the group, the group has power over that member. By power we mean here the ability to produce real change in opinions and attitudes and not simply change in overt behavior which can also be produced by means of overt threat. If a person is unable to leave a group because of restraints from the outside, the group can then use threats to change overt behavior. Covert changes in opinions and attitudes, however, can only be produced by a group by virtue of forces acting on the member to remain in the group. Clearly the maximum force which the group can successfully induce on a member counter to his own forces can not be greater than the sum of the forces acting on that member to remain in the group. The greater the resultant force to remain in the group, the more effective will be the attempts to influence the member.

This hypothesis is corroborated by two separate studies. Festinger, Schachter and Back (4) investigated the relationship between the cohesiveness of social groups in a housing project (how attractive the group was for its members) and how effectively a group standard relevant to the functioning of the group was maintained. A correlation of .72 was obtained between these two vari-

ables. In other words, the greater the attractiveness of the group for the members, the greater was the amount of influence which the group could successfully exert on its members with the result that there existed greater conformity in attitudes and behavior in the more cohesive groups.

Back (1) did a laboratory experiment specifically designed to test this hypothesis. By means of plausible instructions to the subjects he experimentally created groups of high and low cohesiveness, that is, conditions in which the members were strongly attracted to the group and those in which the attraction to the group was relatively weak. The subjects, starting with different interpretations of the same material, were given an opportunity to discuss the matter. Irrespective of the source of the attraction to the group (Back used three different types of attraction in both high and low cohesive conditions) the subjects in the high cohesive groups influenced each other's opinions more than the subjects in the low cohesive groups. In short, the greater the degree of attraction to the group, the greater the amount of influence actually accomplished.

Hypothesis 3c: *The amount of change in opinion resulting from receiving a communication concerning "item x" will decrease with increase in the degree to which the opinions and attitudes involved are anchored in other group memberships or serve important need satisfying functions for the person.*

If the opinion that a person has formed on some issue is supported in some other group than the one which is at present attempting to influence him, he will be more resistant to the attempted influence. Other sources of resistance to being influenced undoubtedly come from personality factors, ego needs and the like.

Specific evidence supporting this hy-

pothesis is rather fragmentary. In the study of social groups in a housing project by Festinger, Schachter and Back (4), the residents were asked whether their social life was mainly outside the project or not. Of those who conformed to the standards of their social groups within the project about 85 per cent reported that their social life was centered mainly within the project. Less than 50 per cent of those who did not conform to the standards of the project social group, however, reported that their social life was centered mainly in the project. It is likely that they were able to resist the influences from within the project when their opinions and attitudes were supported in outside groups.

The experiments by Schachter (8) and by Festinger and Thibaut (5) used the same discussion problem in slightly different situations. In the former experiment subjects identified themselves and verbally supported their opinions in face-to-face discussion. In the latter experiment the subjects were anonymous, communicating only by written messages on which the sender of the message was not identified. Under these latter conditions many more changes in opinion were observed than under the open verbal discussion situation even though less time was spent in discussion when they wrote notes. This difference in amount of change in opinion is probably due to the ego defensive reactions aroused by openly committing oneself and supporting one's opinions in a face-to-face group.

(4) Determinants of change in relationship among members:

Hypothesis 4a: *The tendency to change the composition of the psychological group (pushing members out of the group) increases as the perceived discrepancy in opinion increases.*

We have already discussed two of the responses which members of groups

make to pressures toward uniformity, namely, attempting to influence others and being more ready to be influenced. There is still a third response which serves to move toward uniformity. By rejecting those whose opinions diverge from the group and thus redefining who is and who is not in the psychological group, uniformity can be accomplished. The greater the discrepancy between a person's opinion and the opinion of another, the stronger are the tendencies to exclude the other person from the psychological group.

There is evidence that members of groups do tend to reject those whose opinions are divergent. In the study of social groups within a housing project Festinger, Schachter and Back (4) found that those who did not conform to the standards of their social group were underchosen on a sociometric test, that is, they mentioned more persons as friends of theirs than they received in return. Schachter (8) did an experiment specifically to test whether or not members of groups would be rejected simply for disagreeing on an issue. Paid participants in the groups voiced divergent or agreeing opinions as instructed. In all groups the paid participant who voiced divergent opinion on an issue was rejected on a postmeeting questionnaire concerning whom they wanted to have remain in the group. The same paid participants, when voicing conforming opinions in other groups, were not rejected.

Hypothesis 4b: *When non-conformity exists, the tendency to change the composition of the psychological group increases as the cohesiveness of the group increases and as the relevance of the issue to the group increases.*

We have previously discussed the increase in forces to communicate with increase in cohesiveness and relevance of issue. Similarly, these two variables affect the tendency to reject persons

from the group for non-conformity. Theoretically we should expect any variable which affected the force to communicate (which stems from pressures toward uniformity) to affect also the tendency to reject non-conformers in a similar manner. In other words, increases in the force to communicate concerning an item will go along with increased tendency to reject persons who disagree concerning that item.

The previously mentioned experiment by Schachter (8) was designed to test this hypothesis by experimentally varying cohesiveness and relevance in club groups. In this experiment the more cohesive groups do reject the non-conformer more than the less cohesive groups and the groups where the issue is relevant reject the non-conformer more than groups where the issue is not very relevant to the group functioning. Those groups where cohesiveness was low and the issue was not very relevant show little, if any, tendency to reject the deviate.

### III. FORCES TO CHANGE ONE'S POSITION IN A GROUP

Another important source of forces to communicate are the forces which act on members of groups to locomote (change their position) in the group, or to move from one group to another. Such forces to locomote may stem from the attractiveness of activities associated with a different position in the group or from the status of that position or the like. Thus a new member of a group may wish to become more central in the group, a member of an organization may wish to rise in the status hierarchy, a member of a business firm may want to be promoted or a member of a minority group may desire acceptance by the majority group. These are all instances of forces to locomote in a social structure.

It is plausible that the existence of a

force acting on a person in a specific direction produces behavior in that direction. Where locomotion in the desired direction is not possible, at least temporarily, there will exist a force to communicate in that direction. The existence of a force in a specific direction will produce behavior in that direction. One such kind of behavior is communication. This hypothesis is not very different from the hypothesis advanced by Lewin (6) to account for the superior recall of interrupted activities.

An experiment by Thibaut (9) tends to corroborate this theoretical analysis. In his experiment he created two groups, one of high status and privileged, the other of low status and under-privileged. These two groups, equated in other respects, functioned together so that the members of the high status group could play an attractive game. The low status group functioned merely as servants. It was clear that forces were acting on the members of the low status group to move into the other group. As the privilege position of the high status group became clearer and clearer the amount of communication from the low status team to the high status group increased. The number of communications from members of the high status group to the low status group correspondingly decreased. When, in some groups, the status and privilege relationship between the two teams was reversed toward the end of the experimental session, thus reducing the forces to locomote into the other group, the number of communications to that other group correspondingly decreased.

Further corroboration is found in a preliminary experiment, mainly methodologically oriented, conducted by Back *et al.* (2). In this experiment new items of information were planted with persons at various levels in the hierarchy of a functioning organization. Data on transmission of each of the

items of information were obtained through cooperators within the organization who were chosen so as to give adequate coverage of all levels and all sections within it. These cooperators recorded all instances of communication that came to their attention. Of seventeen acts of communication recorded in this manner, eleven were directed upwards in the hierarchy, four toward someone on the same level and only two were directed downwards. The existence of forces to move upward in such a hierarchical organization may be taken for granted. The great bulk of the communications recorded went in the same direction as these forces to locomote.

In considering communication among members of differentiated social structures it is important also to take into account restraints against communication.

Infrequent contact in the ordinary course of events tends to erect restraints against communication. It is undoubtedly easier to communicate a given item to a person whom one sees frequently or to a person to whom one has communicated similar items in the past. The structuring of groups into hierarchies, social clusters, or the like, undoubtedly tends to restrict the amount and type of contact between members of certain different parts or levels of the group and also undoubtedly restricts the content of the communication that goes on between such levels in the ordinary course of events. These restrictions erect restraints against certain types of communication.

There are some data which tend to specify some of the restraints against communication which exist. In the study of the communication of a spontaneous rumor in a community by Festinger, Cartwright *et al.* (3) it was found that intimacy of friendship tended to increase ease of communication. Per-

sons with more friends in the project heard the rumor more often than those with only acquaintances. Those who had few friends or acquaintances heard the rumor least often. At the same time this factor of intimacy of friendship was not related to how frequently they relayed the rumor to others. In other words, it was not related to forces to communicate but seemed to function only as a restraint against communicating where friendship did not exist.

There is also some evidence that the mere perception of the existence of a hierarchy sets up restraints against communication between levels. Kelley (7) experimentally created a two-level hierarchy engaging in a problem-solving task during which they could and did communicate within levels and between levels. Control groups were also run with the same task situation but with no status differential involved between the two subgroups. There was more communication between subgroups under these control conditions than where there was a status differential involved.

It seems that, in a hierarchy, there are also restraints against communicating hostility upwards when the hostility is about those on upper levels. In the same experiment by Kelley there was much criticism of the *other group* expressed by both high status and low status members. The proportion of these critical expressions which are directed upward by the low status group is much less, however, than the proportion directed downward by the high status groups.

#### IV. EMOTIONAL EXPRESSION

An important variety of communications undoubtedly results from the existence of an emotional state in the communicator. The existence of joy, anger, hostility and the like seems to produce forces to communicate. It



seems that communications resulting from the existence of an emotional state are consummatory rather than instrumental.

By an instrumental communication we mean one in which the reduction of the force to communicate depends upon the effect of the communication on the recipient. Thus in communication resulting from pressures toward uniformity in a group, the mere fact that a communication is made does not affect the force to communicate. If the effect has been to change the recipient so that he now agrees more closely with the communicator, the force to communicate will be reduced. If the recipient changes in the opposite direction, the force to communicate to him will be increased.

By a consummatory communication we mean one in which the reduction of the force to communicate occurs as a result of the expression and does not depend upon the effect it has on the recipient. Certainly in the case of such communications the reaction of the recipient may introduce new elements into the situation which will affect the force to communicate, but the essence of a consummatory communication is that the simple expression does reduce the force.

Specifically with regard to the communication of hostility and aggression, much has been said regarding its consummatory nature. The psychoanalytic theories of catharsis, in particular, develop the notion that the expression of hostility reduces the emotional state of the person. There has, however, been very little experimental work done on the problem. The previously mentioned experiment by Thibaut in which he created a "privileged-underprivileged" relationship between two equated groups has some data on the point. There is evidence that those members of the "underprivileged" groups who expressed their hostility toward the "privileged"

groups showed less residual hostility toward them in post-experimental questionnaires. There is, however, no control over the reactions of the recipients of the hostile communications nor over the perceptions of the communicators of what these reactions were. An experiment is now in progress which will attempt to clarify some of these relationships with both negative and positive emotional states.

#### V. SUMMARY

A series of interrelated hypotheses has been presented to account for data on informal social communication collected in the course of a number of studies. The data come from field studies and from laboratory experiments specifically designed to test the hypotheses.

Three sources of pressures to communicate have been considered:

1. Communication arising from pressures toward uniformity in a group. Here we considered determinants of magnitude of the force to communicate, choice of recipient for the communication, magnitude of change in recipient and magnitude of tendencies to reject nonconformers.

2. Communications arising from forces to locomote in a social structure. Here we considered communications in the direction of a blocked locomotion and restraints against communication arising in differentiated social structures.

3. Communications arising from the existence of emotional states. In this area data are almost completely lacking. Some theoretical distinctions were made and an experiment which is now in progress in this area was outlined.

#### BIBLIOGRAPHY

1. BACK, K. The exertion of influence through social communication. *J. abn. soc. Psychol.*, 1950 (in press).

2. —, FESTINGER, L., HYMOVITCH, B., KELLEY, H. H., SCHACHTER, S., & THIBAUT, J. The methodological problems of studying rumor transmission. *Human Relations*, 1950 (in press).
3. FESTINGER, L., CARTWRIGHT, D., *et al.* A study of a rumor: its origin and spread. *Human Relations*, 1948, 1, 464-486.
4. —, SCHACHTER, S., & BACK, K. *Social pressures in informal groups: a study of a housing project*. New York: Harper & Bros., 1950.
5. —, & THIBAUT, J. Interpersonal communication in small groups. *J. abn. soc. Psychol.* (in press).
6. LEWIN, K. Formalization and progress in psychology. In *Studies in Topological and Vector Psychology I*, Univ. Ia. Stud. Child Welf., 1940, 16, No. 3.
7. KELLEY, H. H. Communication in experimentally created hierarchies. *Human Relations* (in press).
8. SCHACHTER, S. Deviation, rejection, and communication. *J. abn. soc. Psychol.* (in press).
9. THIBAUT, J. An experimental study of the cohesiveness of underprivileged groups. *Human Relations*, 1950, 3.

[MS. received March 6, 1950]

# DYNAMIC SYSTEMS, PSYCHOLOGICAL FIELDS, AND HYPOTHETICAL CONSTRUCTS

BY DAVID KRECH

*University of California*

In a recent article (3) I offered a series of notes and suggestions toward a unified psychological theory. The theory (or rather the program for a theory) as there outlined postulated the existence of a hypothetical construct which I labelled "Dynamic System." With further reflection I now feel it necessary to revise my original definition of the Dynamic System. The necessity for such a revision derives from some considerations which suggest that hypothetical constructs in psychological theory are incompatible with the concept of "psychological field." Since this conclusion, if sound, would call for a drastic reorientation in much of the current psychological theory-building, it appears to me that the reasons for revising my definition of Dynamic Systems may have more general interest than their relevancy to the specific theory of Dynamic Systems. I therefore want to organize the present discussion around two major topics: (1) The nature and role of hypothetical constructs in psychological theory, and (2) The role of neurology in psychological theory.

## HYPOTHETICAL CONSTRUCTS AND PSYCHOLOGICAL THEORY

As was pointed out in my original article, many recent theoretical speculations in psychology have concerned themselves with hypothetical constructs, and a number of theoreticians are coming to depend heavily upon such constructs in their attempts at model-building. Perhaps the most widely accepted definition of what is meant by a hypothetical construct has been pro-

posed by MacCorquodale and Meehl (4) in their discussion of the distinction which can and should be made between a hypothetical construct and an intervening variable. Following their suggestions I pointed out in my original article that, as I intended to use the term,

"A hypothetical construct refers to a postulated actually existing structure which might eventually be described by direct experimentation. As such it is centrally located, and in the sequential history of stimulus-response the functions of the hypothetical construct intervene between stimulus and response. A hypothetical construct, further, will have properties beyond those functional equations through which it was first described. . . . Because it is assumed that these hypothetical constructs exist, and because of the intrinsic properties that they are assumed to have, the correlations between experimental conditions and results are . . . seen as *necessary* correlations, as inevitable consequences of the functioning of these hypothetical constructs."

Having thus described hypothetical constructs I then proceeded to define Dynamic Systems as one instance of a hypothetical construct, but in this specific definition I failed to meet the general requirements I had previously set up for a hypothetical construct. As I have already indicated, my confusion would merit little attention if it betrayed merely my own lack of clear thinking, but it is because I believe that my confusion was, in part, a reflection of a generally accepted position among many psychologists that I want to discuss this problem in some detail.

In my attempt at a definition of a Dynamic System as a *hypothetical construct* I wrote:

"The Dynamic System is assumed to be isomorphically related to a system of neural traces. Dynamic Systems are the major units of our psychological fields and are the resultants of every neural event which is reflected in any kind of brain event."

The faults in the above definition (and they are major faults) are embodied in the two phrases "isomorphically related to a system of neural events" and "major units of our psychological fields." I shall attempt to show in this connection that (a) genuine hypothetical constructs cannot be isomorphically related to a system of neural events, but must be a system of neural events, and that (b) "psychological field" is not and cannot be a hypothetical construct, but can only be a part of the response-complex in the sequential formulation: Stimulus—Hypothetical Construct—Response.

What are some of the hypothetical constructs which have been suggested for psychological theory? To list them all would be a long and perhaps too depressing a task to undertake, but among some of the more current ones are such constructs as motives, needs, tensions, beliefs, attitudes, cognitive structures, expectations, hypotheses, etc.<sup>1</sup> Add to that pluralistic list my attempt at one unifying hypothetical construct, the Dynamic System. Now, if we are to take

<sup>1</sup> I want to make it clear that I am discussing these concepts only when they are used as hypothetical constructs. What I have to say in the following discussion does not apply to these concepts when they are used as intervening variables in the MacCorquodale and Meehl sense, i.e., as merely names for stimulus-response correlations. In other words, my present paper can have relevance and interest only to those psychologists who want to use hypothetical constructs in their system-building.

these as seriously-intended hypothetical constructs (and there are many theoreticians who so intend), then we must assume that the above list refers to postulated actually existing structures which might eventually be described by direct experimentation and which have crucial functions in the entire economy of behavior. I must therefore ask the very simple questions: Where do these hypothetical constructs exist? What is the first approximation to a guess as to how I can begin my search for a more direct examination of the intrinsic properties you have ascribed to them? Failure to answer these questions adequately can only mean that these hypothetical constructs are forever removed from empirical investigation.

What sort of answers have been given, or can be given, by those psychologists who can be labelled as "psychological psychologists" (as opposed to, for example, "physiological psychologists")? It seems to me that usually the answers have shown, or will show, the same confusion that I have displayed in my definition of Dynamic System. The basic nature of this confusion derives from the fact that once we accept the position of a purely psychological psychology we are forced to place all would-be hypothetical constructs in a sort of never-never land—a domain which is forever inaccessible to scientific inquiry. I want now to indicate briefly my reasons for believing this, and to point to some variants of confusion which derive from this basic inadequacy. I should like to anticipate my argument a bit, however, by specifying that this confusion appears only when the question of using hypothetical constructs is involved.

Ask a purely psychological theoretician who proposes to use hypothetical constructs in his theory-making where his constructs—his tensions, expectations, cognitive structures or needs—are



located, and ask him just what they are and how they can be studied more directly so that we can eventually check on his theorizing, and he will say something which boils down to the statement that these hypothetical constructs are in the psychological field, are psychological processes and can be studied by psychological analysis. Persist in your attempt to pin him down further and ask whether tensions or expectations or needs or cognitive structures are merely short-hand terms for certain behavior patterns and he will, of course, deny the implication and point out that he is using them as hypothetical constructs, *i.e.*, these tensions, expectations, needs, etc., are *determinants* of behavior and not behavior itself. Ask him, then, whether these terms refer to conscious experiences, and again he will be forced to answer in the negative. Again he will insist that these terms refer to the processes which lie behind experience and are not to be identified with experience, and he will remind you further that not all needs or tensions or attitudes have conscious correlates. Ask him, finally, whether tensions or needs or cognitive structures are names for neurological activities, and again he will say no, if he is a purely psychological psychologist. (I will discuss the case of the not-so-pure psychologist in a moment.) What then are tensions and needs and cognitive structures? They are not behavior, they are not conscious experiences, they are not neural events. How can we ever study the intrinsic attributes of these non-behavioral, non-experiential, non-neural actually existing structures? The answer again is: They are psychological processes going on in a psychological field and are amenable to psychological analysis. But this answer, in my opinion, is double-talk when we are dealing with hypothetical constructs. I do not object to such an answer because this

"psychological field" is now unavailable to more direct examination than through the study of stimulus-response correlations. That may be unavoidable and merely reflects the present preliminary stage of our science. But what I do object to is that such an answer does not even offer the slightest guess or hunch or even fantasy of how the scientist is *ever* to get beyond the study of correlations between the immediate data of psychology—stimulus and response. This situation, as I will attempt to show presently, is not a fatal one so long as we keep our theorizing on a descriptive level, or an "operational" level (in the narrow sense of the term), or on a correlational level. And there are many psychologists, of course, who intend to do just that. With them we have no quarrel at the present time. But this position of purely psychological psychology is an impossible one to tolerate when we attempt to use hypothetical constructs.

But not all purely psychological psychologists are consistent in their answers to the questions of Where and What. There are those who, in reaction to the difficulties of maintaining a purely psychological position simultaneously with an attempt at using hypothetical constructs, show a bewildering inconsistency in their use of terms. That is, they sometimes define tensions, needs, cognitive structures and the like as hypothetical constructs, and sometimes as conscious experience or as behavior. Thus, for example, they will make statements which can be summarized as follows: "The individual's behavior is determined by his psychological field. His psychological field is determined by his perceptions, needs, tensions, cognitive structures, etc. His perceptions, needs, tensions, cognitive structures, etc., are determined, in turn, by his psychological field." This amounts to saying that (a) the "psy-

chological field" as a hypothetical construct, is a cause of behavior, (b) that "psychological field" is a resultant of other hypothetical constructs—perceptions, needs, etc., (c) that perceptions, needs, etc., are caused by the psychological field. If it is argued that this is a misreading of the position, that all that is meant by the original set of statements is that attitudes and tensions and beliefs and needs are merely the units of the psychological field and that all of these units taken together compose the psychological field, then, I believe, we are still badly off. Either the term "psychological field" refers to the sum total of our conscious experiences at any given moment, or it refers to that *plus* other non-conscious psychological events, or it refers *only* to non-conscious psychological events. If the first interpretation is correct, then the psychological field is merely a name for a specific category of behavior (*i.e.*, "conscious experience") and as such it is not a hypothetical construct which is to be used to explain itself. If the second or third interpretations are taken, then we have, of course, the identical problem with which we confronted the psychological psychologists in the first place—Where and What are these non-conscious psychological events? However, it is my opinion that many of us have shifted from one interpretation to the other, and without realizing the consequences we have been playing a sort of innocent shell-game with the hypothetical construct here at one moment and gone another.

But, the objection may be raised, there are no such purely psychological psychologists as I have been attacking. The point will be made that I have been attacking a S— Man. I have two answers to that. In the first place I would deny the allegation and assert that we have among us quite a number of purely psychological psychologists.

Some of my best friends belong in that category. My second point would be that even the impure psychological psychologists—those who are willing to grant the possibility that neurological events may have something to do with behavior—are caught in exactly the same difficulties of attempting to localize hypothetical constructs in an unknown and an *unknowable* land. The only difference between them and the purists among the psychologists is this: Where the purist attempts to place his hypothetical constructs in a never-never land without further qualifications, the impure one does so but with an almost shamefaced yearning for an alliance with the forbidden land of neurology. The theoreticians who fall into this latter category make much use of such terms as "psychoneural (or neuropsychic) processes," "coordinated with neural events" and "isomorphically related to a system of neural events." Just what is a need or a tension or a cognitive structure when it is described as a *neuropsychic* event? Such a description, it seems to me, can only mean that a need is a neurological event *in addition to* a purely psychological event. Why are we not content with defining a need or a tension or a cognitive structure as a neural event and letting it go at that? Obviously we have been reluctant to do so because we have felt that needs or tensions are *something more* than neural events. But what is this "something more"? I suspect that the reasons behind our feeling that "something more" is involved are of two kinds. In the first place, as I have already indicated, we sometimes think of a need as behavior or conscious experience, and sometimes as a hypothetical construct. As behavior or conscious experience it is obvious that a need is *not* a neural event and so we must add "something more." In the second place, when we do think of a need as a hypo-

thetical construct which lies behind the experience or behavior, we operate on the implicit assumption that a hypothetical construct must itself contain the flavor or attributes of the events it controls. And, of course, the attributes of a purely neurological event do not include the attributes which characterize the experience of need. Therefore, again, we feel that we have to add "something more" than the purely neurological to our would-be hypothetical construct.

The assumption that a hypothetical construct must have the qualities of the events it determines is, however, a wholly unnecessary one and, indeed, tends to diminish the scientific usefulness of hypothetical constructs. Just as the genes which determine the red eye color of the fruitfly do not have any kind of "redness" about them, so is it not necessary that the hypothetical construct of "need" have any of the qualities which we experience when we say we "feel a need" or when we say that that organism is behaving in a "needful fashion." We must always remember that hypothetical constructs have their intrinsic properties, and it is just because they have such properties that they can be so powerful in predicting *new* data. Otherwise they serve no useful purpose, for if we insist on limiting the description of the properties of the hypothetical constructs to those we can observe by direct description of the phenotypic data, then we have succeeded in merely cluttering up a perfectly useful and straightforward descriptive process. It is because the almost-neurological psychologist persists in being a purely psychological psychologist in his actual definitions, that he is no better off than the purist.

I must admit that upon rereading the above paragraphs I am left with an uneasy feeling that I have missed the point of most of the theoretical discussions

and system-buildings of the last fifteen or twenty years. The theme of many of those discussions was that only a psychological analysis of behavior is a fruitful and theoretically mature analysis for psychologists. Perhaps I *have* missed the point, but I prefer to believe that the difficulty lies in the fact that we have never examined very carefully the implications of what is involved in a purely psychological approach to psychology. It now seems to me that a purely psychological approach to psychology has had, and still has, tremendous values—but only for one aspect of the study of psychology—the descriptive one. For system-building which uses hypothetical constructs it is completely inadequate. The purely psychological approach permitted us freedom of operation. It gave us autonomy from the restrictions of an inadequate behaviorism and an equally inadequate neurology. One of the effects of such freedom was a significant improvement in our descriptions of behavior and experience and a permissiveness in coining and using new descriptive terms for new observations. Such descriptions and such correlational analyses are still necessary and are still the proper concern of the psychologist. As long as we remain on that level and do not venture beyond the use of intervening variables in our systemization of our correlations, the purely psychological orientation may be helpful. But the moment we introduce hypothetical constructs into our theory building, then the purely psychological approach becomes untenable.<sup>2</sup> I have argued that

<sup>2</sup> The apparent paradox appears that the purely psychological approach becomes most useful to the most "tough-minded" of the psychologists (those who would remain on the descriptive or correlational level) and least useful to those who would venture into the fantasies of hypothetical construct invention. The reasons for this are, I believe, obvious to the reader by now.

it is untenable because it makes forever impossible any attempt to approach the study of our hypothetical constructs in any more direct manner than through the examinations of the original stimulus-response correlations. This is so, I must repeat, because the psychological position places hypothetical constructs in a domain which, *by definition*, is forever removed from any direct observation (for that domain, it will be remembered, is neither behavioral, experiential or neurological).

The conclusion seems inescapable that a "psychological field" or a "life-space" cannot itself be a hypothetical construct nor can it provide us with the substrate in which we can profitably place our hypothetical constructs.

#### DYNAMIC SYSTEMS AS NEUROLOGICAL EVENTS

Where, then, can we place our hypothetical constructs and what can their nature be? The answer I have come to, on the basis of all of the above considerations, is a simple one and one which is not at all new. It seems to me that the most fruitful thing to do would be to take the plunge and announce that henceforth our hypothetical constructs (through the use of which we hope to understand all behavior and experience) are to be conceived of as molar neurological events—that and nothing more. Such a step amounts to accepting the universe, and such a step may help us to avoid some of the confusions, esotericisms and circular reasonings that we all have been guilty of in times past. Once having made such a step we would then be in a position to manipulate hypothetical constructs, to have fantasies about the intrinsic attributes of these hypothetical constructs and, on the basis of such hunches, to look for new relationships among the primary data of psychology. And what is most significant, such a

step will permit us at least to entertain the hope of eventually being able to study our hypothetical constructs more directly than through guess and hunch.

As I have already said, this position is not at all new. There are some psychologists who have always maintained this position. Thus, for example, Köhler has fairly consistently attempted to localize his hypothetical constructs in the brain tissues of the organism, and his work has already yielded the promise that we may eventually be able to examine his speculations concerning the attributes of his hypothetical constructs by direct experimentation with brain phenomena. Recently, too, Hebb (2) has provided us with an interesting and provocative attempt at a theory of behavior wherein almost all the hypothetical constructs are neurological in character. And, even more recently, support for this position has come from an entirely new and unexpected quarter. I refer to Tolman's (5) recent statement on the need for neurological brain models if psychological theory is to advance. This plea of Tolman's was so unexpected that it even astonished Tolman himself. In discussing a recent symposium on the interrelationships between perception and personality Tolman writes:

"My one plea in connection with these papers would be for a more explicit statement of the neurological brain models which are implied. I make such a statement with surprise because for many years I have objected to what I conceived to be premature neurologizing. It seemed so obvious that psychology was handicapped and led astray by the narrow neurological concepts which it took over uncritically from, say, the physiologist's account of decorticate frogs. And so, for a long time, I have feared any attempt to bolster psychology by trying to base it



on the as yet unobserved or postulated facts of neurology. And I still have qualms about present-day neurology, even if it be called 'Cybernetics'" (5).

The reason Tolman is willing to suppress his qualms is, obviously, not that he believes that neurology now really has come of age and has all the necessary facts which it lacked a number of years ago, but rather that he has realized that *if he is to work with hypothetical constructs* he must define his constructs neurologically—whether neurology is ready for us or not.

I have expressed the hope, above, that by adopting the neurological position we may avoid the confusions, esotericisms and circular reasonings of which we have all been guilty in times past. The retort may be made that neurology being what it is, and psychologists being what they are (especially theoretical psychologists) we shall only be substituting one set of confusions and esotericisms and circular reasonings for another. That, of course, may be. But these latter confusions and errors will be "good errors"; that is, they will be errors of execution but not errors of fundamental approach. I believe it was Erich Fromm who reminded us that the difference between the great scientist and the lesser one is that the former makes productive errors, the latter, sterile ones. And committed, as I am, to the belief that psychology (or any science) will advance only as it incorporates the use of hypothetical constructs in its theorizing, I prefer the real danger of errors of execution to the more serious danger of errors of approach. It is for this reason, if for no other, that I would evaluate an attempt like that of Hebb's as a greater advance for the development of psychology than many previous purely psychological theories.

It is because of all the considerations

presented in this paper, then, that I now want to define my proposed hypothetical construct, Dynamic System, in purely neurological terms. I shall first present a brief definition—or description—and then indicate, equally briefly, some of the implications of the Dynamic System.

The term "Dynamic System" refers to an open system<sup>3</sup> of central neural processes. As such, Dynamic Systems are conceived of as the fundamental units of the brain field. Dynamic Systems are the molar organizations of specified neural events persisting from previous brain activity and of specified neural events deriving from stimulation originating outside the organism or inside the organism. The total brain field, then, can be analyzed into the component Dynamic Systems but no further. The kind and number of Dynamic Systems, the changes which occur within each Dynamic System, and the interaction among Dynamic Systems determine all experience and behavior.

Now for some of the implications of the above brief statement. First it must be emphasized that the above statement only sets the problem. The first task is to specify the attributes of Dynamic Systems *in neurological terms only*. I hope to present such specifications shortly. Secondly it must also be clear that Dynamic Systems are the *only* hypothetical constructs I shall use. That is, as I have indicated at some length in my original article, I find no need for a hypothetical construct to "take care of" so-called motivational data, and another one to take care of

<sup>3</sup> I use the term "open system" as opposed to "closed system" since in my attempts at more precise model-building I have found it more profitable to use the thermodynamics of open systems rather than the thermodynamics of closed systems. For a discussion of some of the differences involved, the reader is referred to the article by von Bertalanffy (1).

so-called cognitive data. That is what I meant by the above phrase that the total brain field "can be analyzed into the component Dynamic Systems but no further." Finally, one of the major tasks, given the above implications, is to spell out the relationships between the postulated neurological attributes of a Dynamic System and the observable psychological data of experience and behavior. It is this final task which is the peculiar task of the psychologist and which differentiates him from the neurologist as such.

## BIBLIOGRAPHY

1. VON BERTALANFFY, L. The theory of open systems in physics and biology. *Science*, 1950, 111, 23-29.
2. HEBB, D. O. *Organization of behavior*. New York: Wiley, 1949.
3. KRECH, D. Notes toward a psychological theory. *J. Personal.*, 1949, 18, 66-87.
4. MACCORQUODALE, K., & MEEHL, P. E. On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.*, 1948, 55, 95-107.
5. TOLMAN, E. C. Discussion. *J. Personal.*, 1949, 18, 48-50.

[MS. received March 17, 1950]

# A QUANTITATIVE DERIVATION OF LATENT LEARNING

BY JAMES DEESE

*The Johns Hopkins University*

A short time ago, in this journal, Seward (3) presented a theoretical derivation of latent learning. The purpose of the present paper is to present a similar derivation. However, in the present case, the derivation is somewhat more general, predicting the results of the experiment of which latent learning is simply a special case. Furthermore, the present derivation has the advantage of predicting the amount of latent learning expected on the basis of making one or another particular assumption. In this way, it allows something of a test of the particular assumption involved. This derivation also has the advantage of being much more simple; it demands a minimum of assumptions and definitions.

If we accept, as a first approximation, Hull's (2) formulation of the way in which habit grows with the number of reinforcements, we may write

$$\frac{dH}{dN} = k(M - H). \quad (1)$$

$H$  is habit strength;  $M$  is the limit which habit strength may take under a particular set of conditions;  $k$  is an arbitrary constant; and  $N$  is the number of reinforcements.<sup>1</sup>

Now, if we assume

$$M = f(D), \quad (2)$$

where  $D$  is drive strength, the relation being such that where

$$D_1 < D_2,$$

<sup>1</sup>Of course,  $N$  can only take discrete and discontinuous values. However, the author feels that he can justify writing this equation by pointing out that  $N$  is just another measure of  $t$ , time in the presence of a reinforcing agent.

then

$$M_1 < M_2.$$

This assumption is implicit in the idea that the amount of drive reduction per unit reinforcement is larger where the absolute level of drive is larger.

The integration of equation (1) yields

$$H = M + Ce^{-kN}. \quad (3)$$

Evaluating the constant  $C$ , we set

$$N = 0,$$

at which point

$$H = 0.$$

Thus

$$C = -M.$$

Substituting the value for  $C$  into (3) we have

$$H = M - Me^{-kN}. \quad (4)$$

Thus, if

$$M = f(D),$$

and

$$D_1 < D_2,$$

then

$$M_1 < M_2.$$

Therefore, for any given value of  $N$

$$H_1 < H_2.$$

At this point it is necessary to introduce another approximation suggested from Hull (2).

$$R = f(D) \cdot f(H). \quad (5)$$

where  $R$  is response tendency (reaction potential in Hull's language). It appears from this equation that, for a particular response to occur, some level of drive strength as well as some level of habit strength must be present.

By defining measurements appropriately, we may write

$$R = D \cdot H. \quad (6)$$

That this may not hold absolutely does not vitiate the derivation.

Now, from equation (4), where  $N$  approaches infinity,  $H$  will approach  $M$  as a limiting value. From this it appears that if we choose a trial,  $k$ , such that  $N$  is very large, we may write (6) as

$$R_k = D \cdot M. \quad (7)$$

Let us suppose that we have reinforced a very large number of times two animals for performing a given response. Also, let us suppose that in one case the drive ( $D$ ) has been large, and in the other case that the drive has been small.

Thus, on trial  $k$  for animal 1 we may write

$$R_{k1} = D_1 \cdot M_1.$$

On trial  $k$  for animal 2 we may write

$$R_{k2} = D_2 \cdot M_2.$$

$R_k$  in animal 2 will be larger than  $R_k$  in animal 1 because both  $D_2$  and  $M_2$  are larger than  $D_1$  and  $M_1$ .

If between trials  $k$  and  $i$  we interchange the drive level of our two animals we will have the following set

of equations for  $R$  in animals 1 and 2 on trials  $k$  and  $i$ . Remember that this defines the state of affairs at the beginning of trial  $i$ , so that the reinforcement from trial  $i$  has not yet had a chance to affect things.

$$R_{k1} = D_1 \cdot M_1,$$

$$R_{k2} = D_2 \cdot M_2,$$

$$R_{i1} = D_2 \cdot M_1,$$

$$R_{i2} = D_1 \cdot M_2.$$

From these equations it follows that

$$R_{i1} > R_{k1}.$$

$$R_{i1} < R_{k2}.$$

$$R_{i2} > R_{k1}.$$

$$R_{i2} < R_{k2}.$$

This means that on trial  $i$ , animal 1 will show more reaction tendency than he did on trial  $k$ , but less than animal 2 showed on trial  $k$ . Animal 2 on trial  $i$  will show less reaction tendency than he did on trial  $k$  and more than animal 1 showed on trial  $k$ . These relationships may be seen in Fig. 1.

Thus, in general, the expected results in a latent learning experiment are predicted. A group of animals trained under low drive would improve suddenly in performance if its drive were increased. In addition we may predict what happens when drive in a

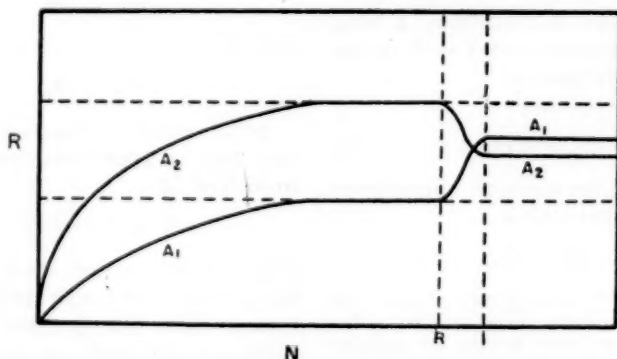


FIG. 1. Showing the change in  $R$ , response tendency, between trials  $k$  and  $i$  on the assumption that  $M$ , the limit of habit strength, is a function of drive.



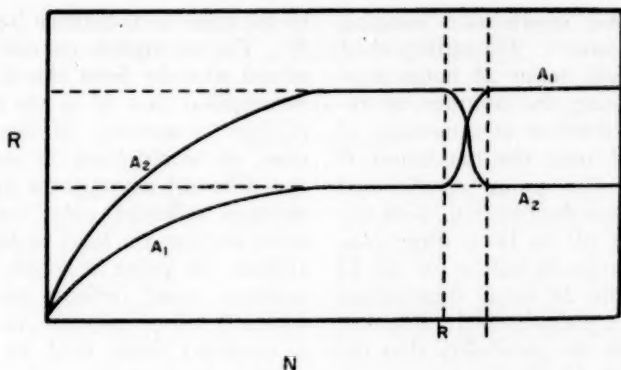


FIG. 2. Showing the change in  $R$ , response tendency, between trials  $k$  and  $i$  on the assumption that  $M$ , the limit of habit strength, is not a function of drive.

second group of animals is decreased in strength.

However, we may change these predicted results by changing one assumed relationship. Let us assume

$$M \neq f(D).$$

This means that  $M$ , the limit of habit strength, is independent of the level of drive. Thus the amount of habit strength acquired per unit reinforcement is also independent of the level of drive. If two animals are trained under two different drives, then their habit strengths would be the same at any given value of  $N$ .

$$H_1 = H_2 = H_c.$$

If  $N$  is allowed to become very large then

$$M_1 = M_2 = M_c.$$

However, at any trial,  $k$ , the response tendencies of our two animals would differ. This would be, of course, because they are operating under different drives.

Suppose again, that between trials  $k$  and  $i$  we interchange drives. We give animal 2 the drive had by animal 1 on trial  $k$ , and we give animal 1 the drive had by animal 2 on trial  $k$ . We

again have four equations for reaction tendency.

$$R_{k1} = D_1 \cdot M_c.$$

$$R_{k2} = D_2 \cdot M_c.$$

$$R_{i1} = D_2 \cdot M_c.$$

$$R_{i2} = D_1 \cdot M_c.$$

From these equations it follows that

$$R_{i1} > R_{k1}.$$

$$R_{i1} = R_{k2}.$$

$$R_{i2} = R_{k1}.$$

$$R_{i2} < R_{k2}.$$

This would mean that the response tendencies of the animals would be interchanged exactly if we interchanged drive. This is shown in Fig. 2. These results are much closer to the theory of latent learning advanced by Tolman's group.

It may be pointed out that Finan (1) has already shown that habit strength is a function of the drive under which it is acquired, and that, by implication,  $M$  is a function of the drive level. This would mean, of course, that the first set of predictions concerning  $R_i$  and  $R_k$  in animals 1 and 2 was correct and that the second set of predictions was wrong. However, there are many problems present in Finan's data. He conditioned animals to press a lever under various

levels of food deprivation ranging from 1 to 48 hours. He extinguished all his animals under 24 hours deprivation. Using the number of responses in extinction as a measure of  $R$ , he found that the maximum  $R$  occurred in the group conditioned under 12 hours deprivation, with the curve falling off in both directions. However, the mean values for the 12 hours and the 24 hours deprivation groups were not significantly different. This suggests the possibility that the results could be deduced from an assumption of generalization of drive stimuli. This hypothesis seems particularly tenable if the distributions of individual scores presented by Finan are examined. At any rate, the problem is by no means settled by Finan's data.

In the present instance a test of these two alternative assumptions is possible. We must only agree upon a good measure of  $R$ . Errors in a complex maze are probably not a good measure of a simple  $R$ . We suggest log latent time as an appropriate measure. It is necessary merely to assume that

$$R = \frac{a}{\log t_L}$$

Any other simple monotonic relationship between log latent time and  $R$  would do; however, this particular assumption simplifies the graphic interpretation. To test these two alternative derivations of latent learning, which differ only in one assumption, it is necessary to train animals to the point at which the asymptote of  $H$  is reached. This can be determined by finding the point at which  $R$  fails

to increase with further increases in  $N$ . The asymptote cannot be determined directly from the data if the assumption that  $M$  is not a function of drive is correct. If this were the case, we would have to assume that the different asymptotes in our two animals reflected only the level of drive and not the level of  $M$ . Nevertheless, the point at which  $R$  fails to increase would indicate the point at which  $H$  fails to increase, because  $D$  is a constant from trial to trial. It would be necessary only to change drive levels in our two animals and observe the  $R$  on the succeeding trial. The author has performed a preliminary experiment upon a small group of animals; thus far it appears that the assumption that  $M$  is not a function of drive fits the data, though the results are not statistically significant. We are now engaged in a larger experiment designed to explore this point. The correctness of one or another of these assumptions, of course, has widespread implication for behavior theory. The results of the above suggested experiment should generate many more experiments. In other words, the idea is heuristic.

#### REFERENCES

1. FINAN, J. L. Quantitative studies in motivation. I. Strength of conditioning in rats under varying degrees of hunger. *J. comp. Psychol.*, 1940, **29**, 119-134.
2. HULL, C. L. *The principles of behavior*. New York: D. Appleton-Century, 1943.
3. SEWARD, J. P. A theoretical derivation of latent learning. *PSYCHOL. REV.*, 1947, **54**, 83-98.

[MS. received January 6, 1950]

## AN IDEAL EQUATION DERIVED FOR A CLASS OF FORGETTING CURVES

BY IVAN D. LONDON

*Harvard University*

The attempt to develop on a rational basis mathematically derived, as opposed to empirically fitted, equations which could satisfactorily prescribe the general course of retention of learned material through time has always come up against a certain difficulty. This difficulty lay in the inability of proposed decay functions to account for the leveling off of various curves of forgetting at some value other than  $n = 0$ , where  $n = \%$  retained as measured by the savings method; otherwise stated, in the inability mathematically to account for the persistent residue of memory over a period of time. In this paper an attempt will be made to derive an ideal equation suitable for a large class of forgetting curves—an equation characterized by constants having rational significance and plausibly based.

Instead of utilizing the concept of a trace which is either strengthened with practice or weakened through disuse—the latter contingency connoting progressive loss in retention of learned material—we shall employ a modified concept of the engram. An engram will be thought of as formed of a complex of activated units, each subject to inactivation and hence to more or less temporary removal from participation in maintenance of the integrity of the engram.

The engram, as thus conceived, will be called an engram-complex in order to distinguish it from the simple notion of a trace. Learning of verbal material or motor activity to given criteria may then be thought of as the establishment of an engram-complex,

the percentage of whose activated units is proportional to the percentage of learned material or activity retained. Forgetting would be the progressive step-like impairment of the engram-complex through summing inactivation of its constituent units. The engram-complex, further, may be visualized as a neural network or, perhaps ultimately to advantage, as a distribution of spatially non-contiguous molecules in a special state of excitation within this neural network.

Upon attainment of a given criterion in the learning process let the subsequent rate of engram-inactivation and, hence, of loss in retention be percentage-wise proportional to the ratio  $n$  of the momentary number of activated units to the original number constituting the initial engram-complex at time  $t = 0$ , that is to say, upon cessation of the learning process. Or, more precisely, let the time rate of change in percentage of activated engram-units be proportional to the momentary percentage present. In other words,

$$\frac{dn}{dt} = -an. \quad (1)$$

The solution of this differential equation is

$$n = ce^{-at}, \quad (2)$$

where  $e$  is the transcendental number encountered in the calculus,  $a$  is the constant of proportionality, and  $c$  is the general constant of integration. Given the initial conditions:  $t = 0$ ,  $n = 1$ , we have

$$n = e^{-at}. \quad (3)$$

The fact that curves of forgetting do not descend sharply to an asymptotic  $n = 0$ , as (3) demands, suggests a restorative process in counter-operation. Now we may think of the activated units of the engram-complex as initiating and/or participating in neurological processes which serve to reactivate the inactivated units and thus to reintroduce them into the impaired engram-complex. Partial restoration of the latter's integrity is thereby continually in progress with a concomitant diminution in rate of loss in retention.

Rate of restoration may, therefore, be seen as proportional *both* to the number of activated units present at any given time *and* to the corresponding number of inactivated units now eligible for reinclusion through reactivation. If  $n$  is the proportion of activated engram-units remaining and  $1 - n$  the proportion rendered inactive in the course of time, the expression to be appended to the right-hand side of (1) as a corrective term is

$$bn(1 - n). \quad (4)$$

The value of  $b$ , the constant of proportionality, reflects, of course, the efficiency of the restorative process. Although there is reason to expect the usual pronounced variation through time in this efficiency, we shall regard it as constant in order to derive an uncomplicated *ideal* equation.

The differential equation which was sought is then,

$$\frac{dn}{dt} = -an + bn(1 - n). \quad (5)^1$$

<sup>1</sup> This is a differential equation of the first order and first degree, and its expanded form identifies it as a *Riccati equation*. This equation is not always solvable in terms of elementary functions (9, p. 39). Fortunately (5) turns out to be solvable in terms of an exponential function, and so an awkward infinite series solution is avoided.

The solution of (5) now demands our attention. On separation of variables, we have

$$\int \frac{dn}{bn^2 - (b - a)n} = - \int dt + c. \quad (6)$$

To evaluate the first term, we proceed as follows:

Let  $k = b - a$ . Then,

$$\int \frac{dn}{bn^2 - kn} = \int \frac{dn}{n(bn - k)}. \quad (7)$$

Employing the method of partial fractions, we find (7) to equal

$$\frac{1}{k} \int \frac{bdn}{bn - k} - \frac{1}{k} \int \frac{dn}{n}; \quad (8)$$

which in turn equals, on integration,

$$-\frac{1}{k} \ln \frac{n}{bn - k}. \quad (9)$$

Substituting in (6), we have

$$-\frac{1}{k} \ln \frac{n}{bn - k} = -t + c. \quad (10)$$

Evaluation of  $c$  when  $t = 0$  and  $n = 1$  and substitution in (10) yield

$$\ln \frac{na}{bn - k} = kt; \quad (11)$$

or, in the preferred exponential form, after some algebraic reshuffling,

$$n = \frac{ke^{kt}}{be^{kt} - a}. \quad (12)$$

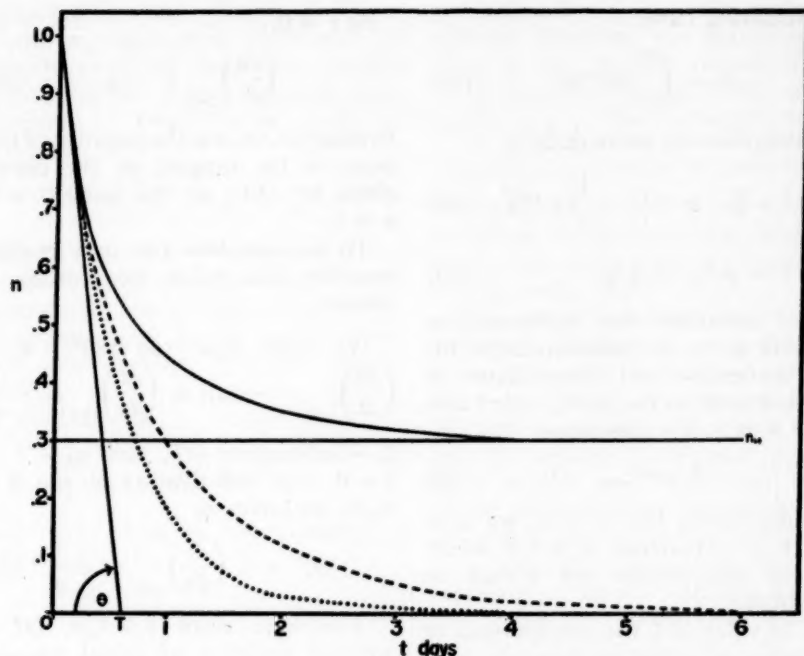
On further simplification, we arrive at the desired ideal equation:

$$n = \frac{b - a}{b - ae^{-(b-a)t}}. \quad (13)^2$$

In contradistinction to those constants which originate in the course of curve-fitting, constants  $a$  and  $b$  have

<sup>2</sup> An infinite series solution, usable for calculation, is  $1/n = 1 + at - a(b-a)t^2/2! + a(b-a)^2t^3/3! - a(b-a)^3t^4/4! + \dots$





Upper curve: equation (13),  $b > a$ . For illustrative purposes, let  $\bar{t} = .6$  days and  $n_{\infty} = .30$ . The corresponding graphed equation, therefore, is  $n = .7/(2.4 - 1.7e^{-.7t})$ . Middle curve: equation (13),  $b < a$ . Here, let  $\bar{t} = .6$  days, but  $b = 1$ . Hence,  $n = .7/(1.7e^{-.7t} - 1)$ . Lower curve: equation (3),  $\bar{t}$  taken as equal to .6 days. Therefore,  $n = e^{-1.7t}$ . See subsequent discussion for evaluation of  $a$  and  $b$ .

rational significance. We proceed to demonstrate this rationality.

We shall first show that  $a$  is the reciprocal of the average duration of engram-unit activation when no restorative process is in operation. We recollect that  $n$  = momentary proportion of activated units in the engram-complex and that  $\frac{dn}{dt}$  = rate of reduction of this proportion. Then, as previously indicated, on exclusion of restorative processes,

$$\frac{dn}{dt} = -an. \quad (14)$$

We saw that the solution of (14) is

$$n = e^{-at} \quad (15)$$

The duration of activation for each individual engram-unit varies and is unpredictable. However, we *can* say that a given proportion of them, equal to  $-\frac{dn}{dt}\Delta t$ , has a temporal span of activation between any  $t$  and corresponding  $t + \Delta t$ . Therefore we may take the average span of activation  $\bar{t}$  as equal to

$$\int_0^{\infty} t \left( -\frac{dn}{dt} \right) dt, \quad (16)$$

where (16) is the expression resulting for the weighted arithmetic mean when  $\Delta t \rightarrow 0$ .

Successive utilization of (14) and (15) enables us to transform (16) into

the following form:

$$\bar{i} = \int_0^{\infty} a t e^{-at} dt. \quad (17)$$

Integration by parts leads to

$$\bar{i} = [-t e^{-at}]_0^{\infty} - \frac{1}{a} [e^{-at}]_0^{\infty} \quad (18)$$

or

$$\bar{i} = n/n + 1/a. \quad (19)$$

To ascertain the indeterminate number  $n/n$ , we differentiate in (18) the numerator and denominator of the first term to the right,  $-t/e^{at}$ , and set  $t = \infty$ . We then have

$$-1/ae^{at}|_{t=\infty} = 0. \quad (20)$$

Substituting (20) in (19), we have  $\bar{i} = 1/a$ . Therefore,  $a = 1/\bar{i}$ , which is the relationship we wished to establish.<sup>3</sup>

The value of  $\bar{i}$ , the average span or duration of initial engram-unit activation, might be thought to be inaccessible. However, interestingly enough, an estimate of its magnitude can be procured graphically.

To see this, we differentiate (13) with respect to  $t$ , getting

$$\frac{dn}{dt} = \frac{-a(b-a)^2 e^{-(b-a)t}}{[b - ae^{-(b-a)t}]^2}. \quad (21)$$

<sup>3</sup> An alternative, though nonrigorous, proof of this relationship proceeds as follows: The rectangular area corresponding to the product of initial maximal  $n$  by average duration of initial engram-unit activation equals total area under curve (15). Therefore, taking  $n = 100\%$  as initial maximum,

$$1 \cdot \bar{i} = \int_0^{\infty} e^{-at} = 1/a.$$

If "half life"  $t_{1/2}$  be taken as the measure of duration of initial engram-unit activation instead of  $\bar{i}$ , then  $t_{1/2}$  can be shown to equal  $(\ln 2)/a$ .

For any engram-unit the probability  $p$  of its initial inactivation between time  $t$  and  $t + \Delta t$  is readily calculable. Substituting (14) in  $p = -\frac{dn}{dt} \Delta t/n$  shows that  $p = a \Delta t = \Delta t/\bar{i}$ .

At  $t = 0$ ,

$$\left(\frac{dn}{dt}\right)_{t=0} = -a. \quad (22)$$

In other words,  $a$  is the negative of the slope of the tangent to the curve, given by (13), at the point  $t = 0$ ,  $n = 1$ .

To discover how one may readily ascertain this value, we proceed as follows:

We note that  $\tan(180^\circ - \theta) = \left(\frac{dn}{dt}\right)_{t=0} = -\tan \theta$ ,  $\left(\frac{dt}{dn}\right)_{t=0} = -\frac{1}{a}$ .

$t_0$  = subtangent  $OT$ , and  $n_0 = 1$  at  $t = 0$ . On substitution in  $\tan \theta = n_0/t_0$ , we arrive at

$$OT = -\left(\frac{dt}{dn}\right)_{t=0} = +\frac{1}{a}. \quad (23)$$

Therefore, since  $\bar{i} = 1/a$ ,  $OT$  = average duration of initial engram-unit activation  $\bar{i}$ . Since for values of  $t$  close to 0, equation (13) is practically linear, there is little difficulty in drawing the tangent corresponding to the point  $t = 0$ ,  $n = 1$ . Hence,  $a = 1/\bar{i}$  can be graphically gauged.<sup>4</sup>

To understand the meaning of the constant  $b$ , our procedure is as follows:

When  $t$  in (13) tends to infinity, the exponential function in the denominator tends to 0 if  $b > a$ . In this case, then,

$$n_{\infty} = \frac{b-a}{b} = 1 - \frac{a}{b}. \quad (24)$$

Thus, given  $a$  and  $b$ , one can measure  $n_{\infty}$ , the degree of integrity of the engram-complex after the lapse of a

<sup>4</sup> Best estimates of  $\bar{i}$  can be anticipated for nonsense syllables with minimal overlearning where the corresponding engram-complex may be expected to enjoy relative isolation from "meaningful" interfering effects, that is to say, from activation and reactivation as a result of supporting liaison with other engram-complexes.

considerable time. Conversely, given  $i$  and  $n_*$ , quantities which may be readily read off from the graphic representation of (13), one can easily compute  $b$ . From (24)  $bn_* = b - a$ . Therefore,

$$b = \frac{a}{1 - n_*} = \frac{1}{i(1 - n_*)}, \quad (25)$$

that is to say,  $b$  = reciprocal of the product of the average span of initial engram-unit activation by the degree of integrity-loss in the engram-complex after the lapse of a considerable time.

Equation (5):

$$\frac{dn}{dt} = -an + bn(1 - n)$$

may with profit be reformulated, if  $b > a$ , as

$$\frac{dn}{dt} = -bn(n - n_*). \quad (26)$$

The transformation proceeds as follows:

Using (25) and simplifying, equation (5) takes on the form:

$$\frac{dn}{dt} = -\frac{an(n - n_*)}{1 - n_*} \quad (27)$$

or, since  $a = 1/i$ ,

$$\frac{dn}{dt} = -\frac{n(n - n_*)}{i(1 - n_*)}. \quad (28)$$

In view of (25), equation (28) may finally be expressed as

$$\frac{dn}{dt} = -bn(n - n_*). \quad (29)$$

Since  $n_*$  corresponds to an asymptotic minimum,  $n_*$  is that particular value of  $n$  which causes the right-hand member of (5) to vanish.<sup>5</sup> The ap-

pearance of  $n - n_*$  as a factor in (29) is, therefore, not surprising in view of the minimal nature of  $n_*$ . This, moreover, would follow as an application of the well-known factor theorem: If  $\frac{dn}{dt}$  is equal to a polynomial

in the single variable  $n$  and if  $n$  approaches a limiting value  $n_*$  with the lapse of time, then  $n - n_*$  is a factor of that polynomial.<sup>6</sup>

We can now restate, by way of summary, the rate of forgetting for non-overlearned material or activity in a particularly instructive form. Because  $n_*$  is the value which  $n$  approaches as time proceeds,  $n_*$  can be aptly termed the equilibrium value or the value at which the corresponding engram-system is at equilibrium. Since equation (5) has already been expressed in the form of (29), we can now state the general law for the engram-complexes that we have been considering: *The rate of engram-impairment at any given moment is proportional to the degree of engram activation obtaining and the departure of the engram-complex from the state of equilibrium.*

If  $b < a$ , equation (13) represents ideally another class of forgetting curves and can be conveniently written as

$$n = \frac{a - b}{ae^{(a-b)t} - b}, \quad (30)$$

which form readily shows that the class of curves represented by (30) is asymptotic to  $n = 0$ , that is to say, after the lapse of a considerable time there is no demonstrable residual memory ( $n_* = 0$ ). In this case, the constant  $b$  loses rationality and its

<sup>5</sup> This fact could have been used to derive (29) from (5) by an alternative method. Subtract  $0 = -an_* + bn_*(1 - n_*)$  from (5), utilizing (25) to simplify resulting equation.

<sup>6</sup> More familiarly stated, if  $\frac{dn}{dt} = f(n)$  vanishes when  $n = n_*$ , then  $f(n)$  is exactly divisible by  $n - n_*$ , and conversely (1, p. 170).

value must be determined by approximation. To facilitate the approximative procedure, let  $m = 1 - b/a$ . Then equation (30) can be transformed into the convenient form:

$$e^{amt} - \frac{(1-n)}{n} m - 1 = 0. \quad (31)$$

Given an estimate of  $a$  and a representative point:  $(t, n)$ ,  $n \neq 0$ ,  $m$  can be approximated with the help of tables of exponential functions. Substitution in  $b = a(1 - m)$  yields the desired value of  $b$ .

Comparison of equation (13) with representative data shows that we have to all appearances succeeded in establishing an ideal equation, mathematically derived and based on plausible grounds. Since (13) is an ideal equation, we must expect deviations from ideal predicted values, but not from the general trend of curvilinear descent asymptotically to  $n \neq 0$ . Granting the law of decay as proposed in (1), deviations are to be expected from the crudity of the savings method as a measure of the underlying condition responsible for the savings performance, from the variety of experimental conditions, criteria, and material or activity employed, from interference effects of one kind or other, from the unknown nature of the restorative function and the undoubted variation in its efficiency as manifested, for instance, in sleeping and waking states. In addition, there is reason to believe that much deviation may be traceable to a neglected factor in learning theory—variation of internal body temperature. The writer offers in brief the following theoretical considerations in support of this statement.

The engram-complex which we have been considering was conceived provisionally as a neural network whose units are liable to inactivation and

potential reactivation.<sup>7</sup> This is a workable and conservative conception. However, in spite of the chance of being premature, it is possible to hazard a more fundamental view of the nature of the postulated engram-complex. Instead of a diffuse assemblage of engram-units in the form of neurones and combinations thereof, we may conceive of our engram-complex as a distribution of highly complex, spatially non-contiguous molecules within a neural network, the concentration of these molecules within any given neurone determining to a large degree both its participation and membership therein.

Let these molecules be capable of assuming quasi-isomeric energy-states (10, pp. 54-55). In other words, they are to exhibit at least two distinct energy-levels corresponding to significantly different and relatively stable configurations of the constituent atoms. The higher energy-level is to be associated with excitation, the lower with its relative lack. According to quantum theory, elevation to a higher energy level and reversion to a lower level are to be understood as discontinuous energy processes. With this interpretation in mind, learning may then be thought of as consisting in the discontinuous raising of energy-levels within a neural network, while forgetting would result from the automatic discontinuous lowering of energy-levels.

From quantum theory we know that, while individual molecular re-

<sup>7</sup> Ideally we must regard the inactivated units as still members of the engram-complex, since they are held reactivatable. Actually, it must be recognized that with time we must expect not only increasing inactivation of successive engram-units, but also growing alienation from the engram-complex, that is to say, an increase in the number of engram-units that are no longer even potentially member-units of the engram-complex.



version from a higher to a lower energy state is in principle unpredictable, reversion for the molecular complex as a whole can and should proceed convergent-wise, that is to say, predictably (4, 5). If we assume that the rate of reversion percentage-wise is proportional to the momentary proportion of excited molecules remaining in the engram-complex, then the whole mathematical formalism which the writer has developed can be retained and reapplied. Now, however, instead of average span or duration of initial engram-unit activation, we must speak of average span or duration of initial molecular excitation.

Given this interpretation of the engram-complex, we may proceed to utilize a mathematical relation involving the Boltzmann exponential factor in order to establish a connection between rate of reversion and temperature (10, p. 64). By means of this relation, which is valid in both classical and quantum theory, the writer has derived the following approximate formula:

$$\bar{i} = i_0 e^{-.14\Delta T}, \quad (32)^8$$

where  $i_0$  = average span of molecular excitation at normal body temperature,  $\Delta T$  = difference in degrees C. between elevated and normal body temperatures,  $\bar{i}$  = average span of molecular excitation at elevated body temperature, and  $e$  is the well-known constant encountered in the calculus.

According to (32), increase in  $\Delta T$  is reflected in decrease in  $\bar{i}$ . Since sub-tangent  $OT = \bar{i}$ , a decrease of  $\bar{i}$  entails a steeper tangent at  $t = 0$ ,  $n = 1$  and, hence, a steeper fall in the curve represented by (13). It is therefore readily seen that rate of forgetting should undergo acceleration

with increase of internal body temperature. Since reversion from higher to lower energy levels takes place in the process of learning as well as in the process of forgetting, learning covering a period of time should grow less efficient with a maintained increase of internal body temperature.

What direct evidence there is relating to these conclusions tends to support the predictions (2, 3, 7, 8). Indirect confirmation may be sought in the reduced rate of forgetting in sleep. It is known that generally the internal body temperature of man varies within a mean range of from 1° to 4° F. daily, the lower temperatures coinciding normally with the nocturnal period of sleep (6). Hence we could conclude on the basis of the adduced theory that the decrease in rate of forgetting demonstrable after a period of sleep is due not alone to the undoubted diminution of interference effects,<sup>9</sup> as usually proposed, but also to the decrease of internal body temperature accompanying sleep.

Deviations from the ideal curve are, then, to be sought also in the over-all variation of internal body temperature. In addition, every time a neurone within the diffuse assemblage constituting the engram-complex fires, local heat conditions are changed. These changes of thermal condition are microscopic in spatial extent, to be sure, but significant for maintenance of the integrity of the engram-complex in accordance to any rule.

The compounding of all these factors with those classically proposed and agreed upon serve to make for pronounced deviations from any derived equation that can be managed. Exact quantitative predictions are, therefore, an impossibility. Only

<sup>8</sup>The derivation of this formula has not been included in this paper, since the derivation is somewhat lengthy.

<sup>9</sup>A diminution contributing to increase in efficiency of the restorative process.

semi-qualitative ones can be reasonably anticipated.

The differential equation (5) and its solution (13), which the writer has proposed, attempt to describe respectively the *ideal* process and course of forgetting around which recorded deviations must be expected more or less to cluster. The molecular interpretation of the engram-complex reinforces the ideal nature of these equations and, in the suggestive guise of (32), offers the means for its experimental verification.

#### BIBLIOGRAPHY

1. FINE, H. B. *College algebra*. Boston: Ginn & Co., 1904.
2. FRENCH, J. W. The effect of temperature on the retention of a maze habit in fish. *J. exp. Psychol.*, 1942, **31**, 79-87.
3. HELLMER, L. A. The effect of temperature on the behavior of the white rat. *Amer. J. Psychol.*, 1943, **56**, 408-421.
4. LONDON, I. D. Some consequences for history and psychology of Langmuir's concept of convergence and divergence of phenomena. *PSYCHOL. REV.*, 1946, **53**, 170-188.
5. —. The need for reorientation in psychology in the light of modern physics. *J. gen. Psychol.*, 1949, **40**, 219-228.
6. LUDWIG, G. D. Rest, work, and sleep. *Industr. Hyg. Dig.*, 1949, **13**, No. 11.
7. MILLS, C. A. Temperature dominance over human life. *Science*, 1949, **110**, 267-273.
8. MOORE, K. The effect of controlled temperature changes on the behavior of the white rat. *J. exp. Psychol.*, 1944, **34**, 70-79.
9. MORRIS, M., & BROWN, O. E. *Differential equations*. New York: Prentice-Hall, 1942.
10. SCHRÖDINGER, E. *What is life?* New York: Macmillan, 1945.

[MS. received January 10, 1950]

# SIMPLE QUALITATIVE DISCRIMINATION LEARNING<sup>1</sup>

BY CLARK L. HULL

*Institute of Human Relations, Yale University*

## INTRODUCTION

At the outset of our consideration of the subject of simple qualitative discrimination learning it will be well to clarify our use of certain terms. It is especially important to distinguish simple discrimination learning from simple trial-and-error learning.<sup>2</sup>

The distinction can be made perhaps most effectively on the basis of the stimulus-response relationships involved. Let it be supposed, for example, that each of two stimuli,  $S_1$  and  $S_2$ , has the capacity to evoke a particular reaction,  $R_1$ , as shown in Fig. 1. Because  $S_1$  and  $S_2$  are qualitatively equivalent in so far as the evocation of  $R_1$  is concerned, this relationship may be said to be that of *stimulus equivalence*. Let it be supposed further that when  $R_1$  is evoked by  $S_1$  ( $S_2$  being absent), the situation will be such that reinforcement will follow, but that when  $R_1$  is evoked by  $S_2$  ( $S_1$  being absent) reinforcement will in no case follow. Under these conditions  $S_1 \rightarrow R_1$  will be progressively strengthened by reinforcement and  $S_2 \rightarrow R_1$  will be progressively weakened by experimental extinction until at length  $S_1$  will uniformly evoke  $R_1$  and  $S_2$  never will; this latter constitutes the state of perfect *simple discrimination*.

<sup>1</sup> The writer is indebted to Harry G. Yamaguchi for aid in the computations which led to Figs. 7 and 8, and to Ruth Hays for aid in the preparation of the manuscript and for permission to use Fig. 3, which came from an experiment performed by her.

<sup>2</sup> This distinction has been emphasized by Spence (14, pp. 429-430). The present article is essentially an elaboration of the writer's interpretation of Spence's extension and formalization (14, 15, 16) of Pavlov's analysis (12, pp. 117 ff.) of discrimination learning.

In summarizing the contrast of the two types of learning just considered we may say that they are alike in that they involve the selective strengthening of one (the adaptive) receptor-effector connection, rather than some other (the unadaptive) receptor-effector connection. They are distinguished by the fact that in simple discrimination learning the receptor-effector connection selected differs on the *stimulus side* from the one eliminated, whereas in simple trial-and-error learning the receptor-effector connection selected differs on the *response side* from that which is eliminated. By extending the meaning of the word *selection* a little, we may say that *simple discrimination learning involves primarily stimulus selection, whereas simple trial-and-error learning involves primarily response selection*.

## A CONCRETE EXAMPLE OF SIMPLE DISCRIMINATION LEARNING

In order that the reader may secure an appreciation of the phenomenon the theory of which we are about to

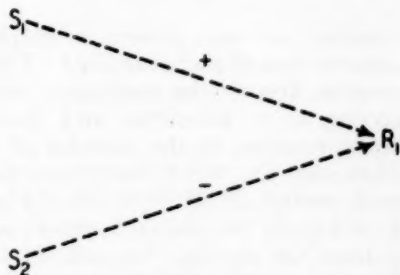


FIG. 1. Diagram showing the type of stimulus-response situation which precipitates simple discrimination learning. Because the reaction potentialities at the outset converge from  $S_1$  and  $S_2$  upon  $R_1$ , this is called the convergent  $S \rightarrow R$  situation.

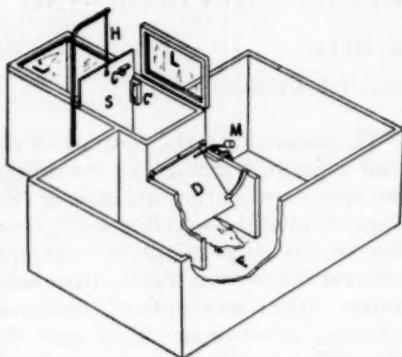


FIG. 2. A drawing of the apparatus utilized in the study of simple white-black discrimination. The albino rat is placed in the chamber beneath the transparent lid marked L, which is shown as closed. When the animal is facing the shutter (S) ready to go into the next chamber the experimenter lifts S somewhat more than enough for the animal easily to pass through into the chamber beneath the lid, L, shown as open, the shutter being suspended in this position by the hooked rod, H. Just as the shutter rises high enough for the animal to pass through, the shoulder C depresses the spring contact C', starting an electric laboratory clock recording time in hundredth seconds. Next, the animal pushes beneath the sloping cardboard door (D) to get the food, F. When the door is raised one inch the microswitch (M) stops the clock which then shows the response time of the subject. The white or black stimuli to be discriminated were placed on the side of the door faced by the rat when in the chamber beneath lid L. Reproduced from Wilcoxon, Hays, and Hull (17).

consider, we now present a simple concrete case of such learning.<sup>3</sup> This consists, first, in the associative connection of a locomotor and door-lifting response to the stimulus of a white card ( $S_1$ ) which constitutes the main portion of the door (D) of Fig. 2, as seen by the animal approaching it from the chamber beneath lid L. The response to white is always followed by food reinforcement. A

<sup>3</sup> The fact that the example given is of quantitative discrimination makes no difference to the reader at the present stage of the theoretical analysis.

curve of this portion of the learning from its beginning approximately up to its asymptote (through days I to V) is shown in Fig. 3. The  $sE_R$  values on the graph were secured by the determination of the median response latency of a group of eight animals, and then the calculation of the equivalent reaction potential by the substitution of the  $sl_R$ 's in the equation,

$$sE_R = 2.57 (sl_R)^{-.614}. \quad (1)$$

Next, the training just mentioned is followed by an irregular alternation between the presentation of the white card ( $S_1$ ) and of a black card ( $S_2$ ). When the door-opening response ( $R_1$ ) follows the presentation of the white card, food reinforcement always follows, but when  $R_1$  follows the presentation of the black card, no food is ever given. This *differential reinforcement*, as it is called, gradually causes a differentiation in the responses to the two stimuli, as shown beginning at day 1 and continuing up through day 18. This represents the peculiarly discriminatory learning. It should be observed that on and around day 15,  $S_2$  (black) evokes a reaction potential fairly close to zero, whereas  $S_1$  (white) evokes a reaction potential of approximately  $3.8 \sigma$ . In this connection it is notable that the reaction potential evoked by  $S_1$  falls very markedly as differential reinforcement progresses for about six days, after which it gradually rises.

We shall now proceed to derive a theoretical account of the empirical phenomena produced by the procedure just presented.

#### THE SPECIAL ROLE OF INCIDENTAL STIMULI ( $S_2$ ) IN DISCRIMINATION LEARNING

The most important type of discrimination learning is based on



stimulus generalization (8, 5, 4). Present available evidence indicates that the gradient of stimulus generalization takes the general form graphically represented in the lower portion of Fig. 4. The phenomenon of stimulus generalization as employed by Pavlov and many other writers (12, 9, 15) is based on a continuous series of potential stimuli and for this reason this potentiality is called a *stimulus continuum*. Such a continuum could be the series of sound (pitch) vibrations varying from high to low, of colors ranging from deep red to orange, of light intensities extending from a strong illumination to one near zero in amount, or of cutaneous vibrations at a constant rate varying from a strong intensity to a weak one. The continuum utilized in the example given above was based on Munsell coated papers ranging from a white having 78.66 per cent of the reflectivity of magnesium oxide to a black having 1.21 per cent.

During learning days I to V, in the experimental situation which produced Fig. 3, it is obvious that in addition to the whiteness of  $S_1$  (i.e., of the generalization continuum), many additional or incidental stimuli presum-

ably become conditioned to  $R_1$ . These include other stimuli which consistently impinge on the organism's sensorium during the repeated reinforcements. Examples of incidentally conditioned stimuli in the situation under consideration are the sound of the click when the sheet-metal shutter (S) is lifted, the lights and shadows of the chamber which the animal passes through from shutter to door (D), any odors that may be consistently present, and the infinitely complex stimuli coming from the animal's own body. We shall call these non-continuum stimulus elements the *incidental* or *static* stimuli. They will be represented as a whole by  $S_3$ . The learned attachment of  $S_3$  to  $R_1$  obviously sets up a separate reaction potential which operates wherever  $S_3$  occurs, and  $S_3$  always accompanies the generalization continuum as used in the experiment. This means that the  $S_3$  stimuli by themselves will evoke  $R_1$  whenever  $S_1$  and  $S_2$  are presented, quite apart from the latter. Thus, so far as  $S_3$  is concerned, this combination of circumstances may give a superficial appearance of 100 per cent generalization.

Since both sets of stimuli,  $S_1$  and

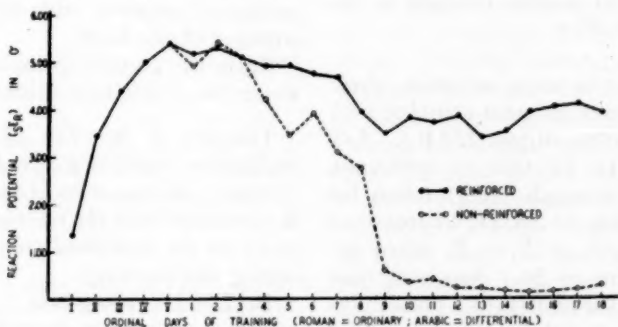


FIG. 3. Graphic representation of the original learning (days I to V) and of the results of the differential reinforcements (days 1 to 18). Note that despite the fact that  $S_1$  is frequently reinforced and never receives a non-reinforcement, it ends with a much lower  $sE_R$  than was displayed on day V. Graph based on data from an unpublished study performed in the writer's laboratory by Ruth Hays.

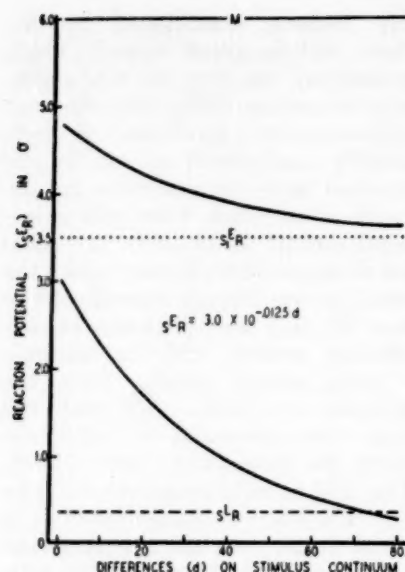


FIG. 4. The lower portion of this figure shows a graphic representation of a theoretical generalization gradient on a qualitative stimulus continuum plotted in terms of j.n.d. differences. Above is shown the same gradient after it has been combined (+) with a reaction potential of  $3.5 \sigma$  (dotted horizontal line) assumed to be caused by incidental stimuli ( $S_4$ ). Note that this summation greatly distorts the generalization gradient (a) by moving it upward leaving its asymptote at the level of  $3.5 \sigma$ , and (b) by greatly reducing its slope. The value of  $M$  is represented by the upper horizontal line at  $6.0 \sigma$ , and that of the reaction threshold by the broken line at  $.355 \sigma$ .

$S_3$ , involve the same response, their reaction potentials must combine (+) with each other (9, pp. 222 ff.). Let us assume (11, IV) that the irrelevant ( $S_4$ ) habit strength ( $s_H R$ ) taken by itself amounts to .58333, whereas the habit strength of  $S_1 \rightarrow R_1$  taken by itself amounts to .50. Assuming that  $M$  equals  $6.0 \sigma$  and (11, IX) multiplying, we find that the reaction potential controlled by  $S_1$  directly is  $.50 \times 6.0$ , or  $3.0 \sigma$ , and that controlled by  $S_3$  is  $.5833 \times 6.0$ , or  $3.5 \sigma$ . Combining this constant incidental reaction po-

tential of  $3.5 \sigma$  with each point of the generalization gradient by means of summational equation 2,

$$s_1 E_{R_1} + s_4 E_{R_1} = s_1 E_{R_1} + s_4 E_{R_1} - \frac{s_1 E_{R_1} \times s_4 E_{R_1}}{M}, \quad (2)$$

we secure the displaced and distorted gradient shown above in Fig. 4.

It is important to compare these two manifestations of the same gradient. The addition of the irrelevant  $s_4 E_R$  of  $3.5 \sigma$  artificially raises the asymptote of the generalization gradient by that amount. At the same time, summation greatly flattens the gradient, though the fractional amount of fall toward the asymptote of each gradient at each increment of  $d$  is constant and exactly equal in both, in the present case approximately 56.23 per cent of the preceding point. However, if the value of  $M$  were reduced from  $6.0 \sigma$  to  $5.0 \sigma$ , the upper gradient would appear much less steep still, i.e., it would be even more distorted, and if  $M$  were reduced to  $3.6 \sigma$  the gradient would be so flat as to be practically horizontal; empirically it would hardly be detectable. At bottom this is because near the asymptote in ordinary learning (3) additional practice adds little to the strength of the habit.

From the preceding considerations, we arrive at our first theorem:

**Theorem 1. A.** *The stimulus generalization gradient as produced on a stimulus continuum by simple learning is summated with the reaction potential based on the incidental stimuli present during that learning.*

**B.** *This summation artificially raises the apparent asymptote of the gradient and flattens the gradient. As the incidental reaction potential approaches the magnitude of  $M$ , the gradient approaches the horizontal.*

Theorem 1, as illustrated by Fig. 4, finds substantial corroboration in the gradients reported by Hovland (6, 7). Incidentally, this may explain why Pavlov's (12, p. 118), Hovland's, and Brown's (1) generalization gradients were approximately horizontal on the first unreinforced trial utilized in the tests for generalization (see Fig. 5, first trial). As a matter of fact, practically the same thing is shown by Fig. 3 at days 1, 2, and 3, by the near equality of the white ( $S_1$ ) and black ( $S_2$ ) values, even though these are each based on eight differential reinforcements.

Next there arises a question of how, in case the incidental  $sE_R$  nearly equals the value of  $M$ , the generalization gradient can ever become manifest. The answer is that its appearance results from the gradual removal of the incidental  $sE_R$  by experimental extinction. Frequently, however,  $S_3 \rightarrow R_1$  cannot be extinguished without the non-reinforced presentation of the stimulus continuum in some form. For example, in the

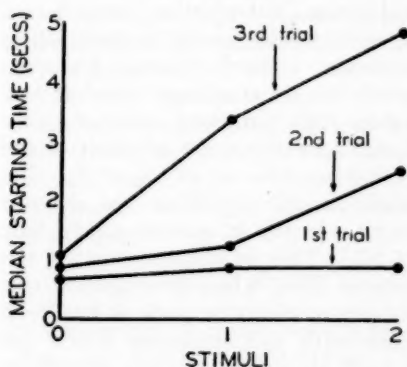


FIG. 5. Graphic representation of the gradual appearance of the stimulus generalization gradient during the process of differential reinforcement. Note that the gradient is here based on response latency and that when so represented the gradient rises, instead of falling as it does when plotted in reaction potentials. Reproduced from Brown (1, p. 218).

present experiment a door is required, and any door must be of some shade. Therefore, in order not to complicate the theoretical picture unduly, let us assume that when  $S_3$  is presented for differential reinforcement, the most favorable value on the stimulus continuum to use with it is the black actually utilized in Fig. 3 as  $S_2$ , because it will presumably evoke a very weak or subthreshold  $sE_R$ . We shall accordingly neglect for the present the  $sE_R$  connected by stimulus generalization to  $S_2$ , and concentrate on the  $sE_R$  based on  $S_3$ .

At the outset of differential reinforcement, the incidental reaction potential of  $S_3 \rightarrow R_1$  presumably is well along toward its asymptote (Fig. 3) from the training of days I to V. This implies that the extinction portion of the differential reinforcement training will start from near zero and presumably (17) will be advancing relatively rapidly. It follows that during differential reinforcement the  $sI_R$  will tend gradually to overtake the  $sE_R$  and the superthreshold portion of the  $s\ddot{E}_R$  will gradually approach zero, i.e., there will be a tendency toward<sup>4</sup>

$$s\ddot{E}_R - sI_R = 0. \quad (3)$$

This means in effect that the effective reaction potential under the control of the incidental stimuli will for most purposes gradually become relatively neutral and unimportant in the determining of overt action. However, it may be added that many stimuli which are present nearly all the time are believed to be neutralized early in life for most responses by the process just described, and therefore do not go through the further neutralization process.<sup>5</sup>

<sup>4</sup>  $s\ddot{E}_R = sE_R - sL_R$ .

<sup>5</sup> This principle of the previous neutralization of incidental stimuli may play an important role in the transfer of training in discrimination learning.

In terms of our equations as applied to Fig. 4, we simply reduce the values of the summated (+) but distorted gradient shown above, by 3.5  $\sigma$  less the value of  $sL_R$ . This is substantially a reversal of the summational procedure (equation 2) which produced it in the first place. The symbolic reversal (withdrawal, or  $\div$ ) is performed for us by equation 4(11).

$$s_1 E_{R_1} = \frac{M(C - s_3 E_{R_1})}{M - s_3 E_{R_1}}, \quad (4)$$

where  $C = s_1 E_{R_1} \div s_3 E_{R_1}$ .

This of course substantially reveals the true gradient in its proper position as represented at the bottom of the figure with which we began our analysis.

Generalizing on these considerations, we arrive at our second theorem:

**Theorem 2.** *As the differential reinforcement trials of the discrimination learning process take place one by one, leading as they do to the practical neutralization of  $S_3 \rightarrow R_1$ , the generalization gradient based on the stimulus continuum must gradually emerge with a progressively steepened gradient, standing in substantially its true form and position at the end of the process except that its level at  $S_2$  must necessarily be at the reaction threshold ( $sL_R$ ).*

While not duplicating the conditions of the theorem, two independent studies, one by Hovland (6) and one by Brown (1) show a clear increase in the steepness of the generalization gradient as the extinction incidental to testing for the presence of the gradient progresses. The Brown results are reproduced as Fig. 5. Note the progressive increase in the slope of the gradients as the ordinal number of trials increases. Incidentally the level rises at the point conditioned, showing that  $sE_R$  is being withdrawn all along the generalization gradient.

To be wholly convincing, of course, such an experiment must remove only the  $S_3 \rightarrow R_1$ , whereas Brown's and Hovland's experiments, since the testing trials were not reinforced, also presumably removed a portion of  $S_1 \rightarrow R_1$  as well.

#### DIFFERENTIAL REINFORCEMENT APPLIED TO $S_1$ (+) AND $S_2$ (-) ON THE STIMULUS CONTINUUM ONLY

With the theoretical elimination of the influence of the incidental reaction potential ( $S_3$ ) by having it reduced in strength to a practical threshold value, we may now consider the effects of differential reinforcement on the stimulus continuum without this complication. This is in a sense the heart of discrimination theory.

It will be recalled that the stimulus continuum of our example was white ( $S_1$ ) through the series of intervening grays down to black ( $S_2$ ), the response ( $R_1$ ) being originally connected to white by simple reinforced association. Now by the principle of stimulus generalization, the reaction potential of  $S_1 \rightarrow R_1$  would extend in diminishing amounts toward black. Unfortunately we do not know to what low values this gradient spontaneously falls, i.e., without the influence of the extinction effects of  $S_2 \rightarrow R_1$ . Indeed, strictly speaking we do not know whether it spontaneously falls at all. This is because, up to the present time, when an empirical test for the gradient is made it has been done with the responses which are evoked to the various test stimuli on the continuum always non-reinforced which would naturally introduce a certain amount of extinction all along the continuum tested for generalization. This has been considered (7, 10, p. 125) to indicate that the parts of the continuum remote from the

point positively reinforced may be more susceptible to the influence of experimental extinction than are those which are closer, and that this differential resistance to extinction may be an indispensable contributing cause producing the gradient.

However this may be, one thing is clear: the conditioning of  $S_1$  to  $R_1$  at one point of a stimulus continuum in fact sets up the potentiality of a generalization gradient which becomes manifest either (a) by differential resistance to extinction (7, 10), or (b) by the removal of a distorting reaction potential attached to incidental stimuli ( $S_3$ ). Until this uncertainty is removed by appropriate experiment the present writer is now ready to gamble that the dominant mechanism in the bringing out of the generalization gradient is the removal of the incidental  $S_3 \rightarrow R_1$  reaction potential, as pointed out above. The existence of incidental stimuli in most such situations is an obvious fact. Moreover, this deduction explains at the same time the steepening of the gradient and the lowering of its asymptote. On the other hand, the hypothesis of differential resistance to experimental extinction is purely *ad hoc*.

The lower gradient on Fig. 6 is accordingly constructed on this hypothesis,  $S_1$  necessarily falling on a  $d$  value of zero. The values of this gradient are calculated by means of equation 5:

$$sE_R = 4.0 \times 10^{-.0125d} \quad (5)$$

In the differential reinforcement process in this case, the  $S_2$  is assumed to fall at  $d = 30$ , which reduces the reaction potential at that point to the reaction threshold, which is taken as .355 (3). Now this reduction in reaction potential is due in the main to the accumulation of conditioned inhibition. Moreover, conditioned in-

hibition generalizes on the stimulus continuum substantially as does reaction potential (11, XI C). The slope of neither gradient is known with precision, but they presumably do not differ very much. We shall accordingly here assume them to be equal.

But in order to calculate the generalization gradient of  $sI_R$ , we must know the amount that is to be withdrawn from the  $sE_R$  at  $d = 30$ . Now this  $sE_R$  value is 1.6868  $\sigma$ , and this withdrawal must be of the  $\pm$  variety which requires the use of equation 4. In this equation, therefore,

$$\begin{aligned} C &= 1.6868 \sigma, \\ sE_{R_1} &= sL_R = .355 \sigma, \\ M &= 6.0 \sigma, \end{aligned}$$

and

$$sE_{R_1} = sI_R.$$

Substituting in equation 4, we have,

$$\begin{aligned} sI_R &= \frac{M(C - sE_{R_1})}{M - sE_{R_1}}, \\ &= \frac{6.0(1.6868 - .355)}{6.0 - .355}, \\ &= \frac{6.0 \times 1.3318}{5.645}, \end{aligned}$$

$$\therefore sI_R = 1.4156.$$

Accordingly the generalization gradient of conditioned inhibition is calculated on the equation,

$$sI_R = 1.4156^{-.0125d}.$$

It is plotted as the shaded double-winged gradient at the bottom of the figure, inverted as it must be when plotted against the negative scale at the left.

Next, the generalized conditioned inhibition must be withdrawn ( $\pm$ ) at each point from the corresponding reaction potential by the repeated use of formula 4. A graphic representation of these differences is shown as a broken line between the two basic



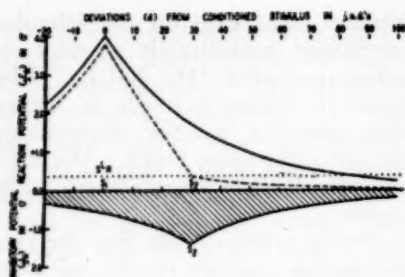


FIG. 6. Graphic representation of the theoretical interaction ( $\pm$ ) of the gradients of the stimulus generalization of reaction potential (upper solid curves) and of conditioned inhibition (lower solid curves) which produce the discrimination gradients (broken-line curves) between. By sighting along the section of this latter line from  $S_1$  to  $S_2$ , it may be seen that this discrimination gradient proper still has a slight curvature remaining from the primary stimulus generalization gradient from which it was derived. Note also that its maximum ( $s\bar{E}_R$ ) lacks a little of the original maximum  $sE_R$ . This same value appears in Fig. 7 at  $d = 30$  (8, p. 25).

gradients. The portion of this difference function falling between  $S_1$  and  $S_2$  will be called the *discrimination gradient*. It is evident from a glance at Fig. 6, that this discriminatory gradient is much more nearly linear than an ordinary stimulus generalization gradient, and that its maximum height, the *net discriminatory potential* ( $s\bar{E}_R$ ), is appreciably less than the original stimulus generalization gradient above it.

Consider now the results of the application of the differential reinforcement procedure to  $S_1$  and  $S_2$ , when the latter stands at  $d = 90$ . In this latter case the  $sE_R$  already stands below the reaction threshold, which means that the maximum  $sI_R$  at that point will be zero and there will be no  $sI_R$  to be withdrawn anywhere from the stimulus generalization gradient. Accordingly, on these assumptions this discrimination gradient will be the stimulus generalization gradient

without change. If  $S_2$  is moved up to a  $d$  of 80 or 70, it is evident to inspection that the maximum amount of  $sI_R$  will be so small that its generalization will not change the stimulus generalization gradient appreciably. However, as  $S_2$  moves up to  $d = 60$  and especially to 50, it is clear that the curvature of the discrimination gradient will grow less, its slope will grow more steep, and its height ( $s\bar{E}_R$ ) will also grow less, ultimately becoming zero when  $d$  is 0. Various theoretical discrimination gradients of this type, as computed from the above assumptions, are represented in Fig. 7. The  $d$  value involved in each is indicated by the lowest level ( $sL_R$ ) reached by each.

Because of the theoretical importance of the relationship, we have also plotted the net discriminatory reaction potential ( $s\bar{E}_R$ ) remaining after complete differential reinforcement as a function of  $d$ . This is reproduced as Fig. 8. There it may be seen that the  $s\bar{E}_R$  begins falling very slowly as  $d$  decreases, and falls progressively more rapidly as  $d$  approaches zero.

Finally it may be stated that the ratios of  $s\bar{E}_R$  to  $d$  were calculated for the values represented in Figs. 7 and 8. They were found to increase progressively as  $d$  approaches zero.

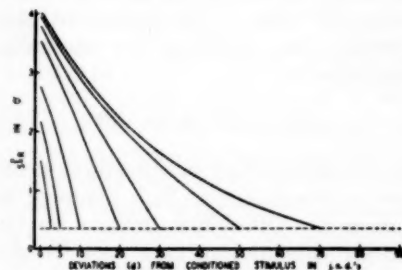


FIG. 7. Graphic representation of the theoretical discrimination gradients resulting from the differential reinforcement of discrimination learning at progressively smaller differences ( $d$ 's) between  $S_1$  and  $S_2$ .

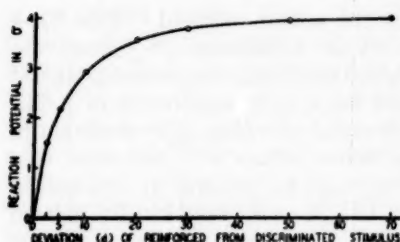


FIG. 8. Graph representing the net discriminatory reaction potential ( $sE_R$ ) as the difference ( $d$ ) between the reinforced and the discriminated stimulus decreases with concurrent differential reinforcement.

Generalizing on the preceding considerations, we arrive at our third theorem:

Theorem 3. A. When differential reinforcement is applied to  $S_{1+}$  and  $S_{2-}$ , with  $d$  practically the entire range of the stimulus continuum, the gradient of  $sE_R$  evocable by the stimulus continuum as a whole presents a curve which is markedly concave upward with an asymptote which is the reaction threshold ( $sL_R$ ).

B. As the range of  $d$  is decreased, the discrimination gradient connecting  $S_1$  and  $S_2$  becomes less concave upward, passing into a slightly convex-upward form after  $d = 30$ .

C. As the range of  $d$  is decreased, the more nearly vertical will tend to be a straight line drawn from the reaction potentials of  $S_1$  and  $S_2$ .

D. As the range of  $d$  is decreased, the smaller will be the net discriminatory reaction potential attached to  $S_1$ ,  $sE_R$  equalling zero when  $d$  equals zero.

E. As the range of  $d$  is decreased, the ratio of the  $sE_R$  at  $S_1$  to  $d$  increases with a positive acceleration.

Up to the present time the extremely simple form of discrimination learning which has been the basis of the preceding analysis has been reported in only three published studies, those by

Frick (2), Grice (4), and Raben (13). So far as reported, none of these studies disagrees noticeably with Theorem 3. Frick (2, p. 119) presents results which tend to substantiate Theorem 3 A, and Raben reports data (13, p. 267) which are in general harmony with 3 D.

We may also take this occasion to consider the learning curves shown in Fig. 3. The curve of the first five days (I to V) is evidently the ordinary growth curve of simple learning (11, IV). With the beginning of differential reinforcement, the distinctive sort of learning characteristic of discrimination begins. The extinction effects ( $sI_R$ ) on  $S_3$  and on  $S_2$  begin to produce a fall on the non-reinforced curve. And since the reinforced curve must also gradually lose the  $S_3 \rightarrow R_1$  component (Fig. 4) at the beginning of the differential reinforcement, this also should fall. In addition, the extinction effect ( $sI_R$ ) on  $S_2$  generalizes to some extent to  $S_1$ , which also contributes to the fall of the reinforced curve, though not so much. But  $S_1$  is reinforced at every response in which it is involved. It follows that the non-reinforced curve falls practically to its reaction threshold as differential reinforcement continues, whereas the reinforced curve remains fairly high. However, soon after the most rapid extinction of  $S_2$  (and  $S_3$ ) has occurred, the continual reinforcement of  $S_1$  apparently causes a small rise in the reinforced curve, presumably because this curve had already become rather close to, though it had not quite reached, its asymptote shortly before.

#### AN EXPERIMENTAL PROBLEM

One of the matters which is likely to arouse a difference of opinion among serious students of behavior is the working hypothesis expressed above,

that the initial empirical horizontal gradient of generalization is mainly due to the summation (+) of the necessarily horizontal gradient of  $s_1E_{R_1}$  with the postulated curved gradient extending from  $s_1E_{R_1}$  to  $s_2E_{R_1}$ . It must be admitted that the evidence in this case is at present far from clear. This situation ought not to continue in a case so fundamental for general behavior theory as this one is. Fortunately the problem would appear to be susceptible to a fairly direct experimental approach. Both time and energy fail the author to perform the task. Partly as a means of further elucidating the present theoretical approach, there are given some suggestions for a method of differentially quantifying the assumed reaction potentials of  $S_1 \rightarrow R_1$  and  $S_2 \rightarrow R_1$ :

Let it be supposed that we wish to determine empirically the reaction potential to sound intensity of  $S_1 \rightarrow R_1$ , uncomplicated by  $S_2 \rightarrow R_1$  values. We would begin by setting up a reinforced connection between a 100-decibel sound and the response to the apparatus such as that shown in Fig. 2, on one hundred albino rats. After this habit had reached its asymptote the median reaction latency ( $sl_R$ ) of the group would be determined and this converted into reaction potential by equation 1 (3). This presumably would be the summation (+) of the reaction potentials evoked by  $S_1$  and  $S_2$ .

Then the 100-decibel sound stimulus component would be physically withdrawn from the stimulus complex, the median response latency noted, and this would be converted into reaction potential by equation 1. This value presumably would be the  $S_2 \rightarrow R_1$  magnitude except for possible interaction effects (11, XII). It should also be appreciably less than that be-

fore the sound removal. This  $S_2 \rightarrow R_1$  value withdrawn (+) from the original total reaction potential should yield the approximate value of  $s_1E_{R_1}$ .

A useful checking approximation to the above values with the same subjects could be secured by extinguishing the  $S_2 \rightarrow R_1$  reaction by massed trials. Substituting this  $\bar{n}$  in an appropriate equation,  $sE_R = f(\bar{n})$  (11, XVII A), the equivalent,  $s_2E_{R_1}$  value could be secured. Then by restoring the sound and extinguishing again, substituting the new  $\bar{n}$  in the equation,  $sE_R = f(\bar{n})$ , the  $s_1E_{R_1}$  value (except for interaction effects) presumably would be secured.

#### SUMMARY

Discrimination learning in simple form is the acquisition of the power of responding differentially to stimuli,  $S_1$  and  $S_2$ , which originally are more or less equivalent to each other in response ( $R_1$ ) evocation. This process of differentiating the response-evocation potentiality of  $S_1$  and  $S_2$  is complicated by certain active stimuli ( $S_3$ ) which are present with both  $S_1$  and  $S_2$ . If the learning process begins by a reinforced association between the stimuli present and  $R_1$ , this necessarily sets up a similar reinforced association between  $S_3$  and  $R_1$  which will function both when  $S_1$  is present and when  $S_2$  is present, giving an appearance of generalization of the stimulus complex even if  $S_1$  and  $S_2$  were quite without generalization capacity.

The primary process which gives rise to the discrimination of the stimulus complex is differential reinforcement. In simple discrimination this consists in reinforcement when the correct stimulus ( $S_1$ ) evokes the reaction, and in non-reinforcement when the incorrect stimulus ( $S_2$ ) evokes the reaction. Since  $S_3 \rightarrow R_1$

is reinforced every time  $S_1 \rightarrow R_1$  occurs, and is non-reinforced every time  $S_2 \rightarrow R_1$  occurs, in the course of time with an equal number of occurrences of  $S_1$  and  $S_2$  both  $s_3E_{R_1}$  and  $s_3I_{R_1}$  tend to approach their asymptotes, which will reduce  $sE_R$  toward its reaction threshold. Moreover, since the rates of both learnings are reduced as the asymptotes are approached, presumably  $S_3$  also loses much of its capacity to acquire not only response  $R_1$  but also similar responses employing the same effectors.

Along with the process of neutralizing  $S_3$  to  $R_1$ , the differential reinforcement gradually builds up  $S_1 \rightarrow R_1$ . But this reaction potential generalizes to  $S_2 \rightarrow R_1$ , which is never reinforced, giving rise to  $s_3I_{R_1}$  which generalizes to  $S_1 \rightarrow R_1$  and reduces the latter appreciably. The net result of this double generalization of  $sE_R$  as  $sI_R$  is the marked loss by  $S_2$  of the power to evoke  $R_1$  and usually the retention of a considerable though reduced power of  $S_1$  to evoke  $R_1$ . This limitation is comparatively slight when the deviation ( $d$ ) between  $S_1$  and  $S_2$  is great, but increases with a positive acceleration as  $d$  approaches zero.

## REFERENCES

1. BROWN, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1943, **33**, 209-226.
2. FRICK, F. C. An analysis of an operant discrimination. *J. Psychol.*, 1948, **26**, 93-123.
3. GLADSTONE, A. I., YAMAGUCHI, H. G., HULL, C. L., & FELSINGER, J. M. Some functional relationships of reaction potential ( $sE_R$ ) and related phenomena. *J. exp. Psychol.*, 1947, **37**, 510-526.
4. GRICE, G. R. Visual discrimination learning with simultaneous and successive presentation of stimuli. *J. comp. physiol. Psychol.*, 1949, **42**, 365-373.
5. GRICE, G. R., & SALTZ, E. The generalization of an instrumental response to stimuli varying in the size dimension (in press).
6. HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125-148.
7. —. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279-291.
8. HULL, C. L. The problem of stimulus equivalence in behavior theory. *PSYCHOL. REV.*, 1939, **46**, 9-30.
9. —. *Principles of behavior*. New York: D. Appleton-Century Co., Inc., 1943.
10. —. The problem of primary stimulus generalization. *PSYCHOL. REV.*, 1947, **54**, 120-134.
11. —. Behavior postulates and corollaries—1949. *PSYCHOL. REV.*, 1950, **57**, 173-180.
12. PAVLOV, I. P. *Conditioned reflexes* (trans. by G. V. Anrep). London: Oxford Univ. Press, 1927.
13. RABEN, M. W. The white rat's discrimination of differences in intensity of illumination measured by a running response. *J. comp. physiol. Psychol.*, 1949, **42**, 254-272.
14. SPENCE, K. W. The nature of discrimination learning in animals. *PSYCHOL. REV.*, 1936, **43**, 427-449.
15. —. The differential response in animals to stimuli varying within a single dimension. *PSYCHOL. REV.*, 1937, **44**, 430-444.
16. —. The basis of solution by chimpanzees of the intermediate size problem. *J. exp. Psychol.*, 1942, **31**, 257-271.
17. WILCOXON, H. C., HAYS, R., & HULL, C. L. A preliminary determination of the functional relationship of effective reaction potential ( $sR_R$ ) and the ordinal number of Vincentized extinction reactions ( $n$ ). *J. exp. Psychol.*, 1950, **40**, 194-199.

[MS. received January 21, 1950]

## A NOTE ON DEPTH PERCEPTION, SIZE CONSTANCY, AND RELATED TOPICS

BY HAROLD SCHLOSBERG

*Brown University*

Recently there has been considerable interest in certain demonstrations by Ames (1, 2). Cantril (5) has done further work with these demonstrations (*Cf. Life*, Jan. 16, 1950, pp. 57-62). Many of these demonstrations indicate that the direct physiological clues, such as disparity and accommodation, are relatively less important in determining depth perception than are factors like knowledge (true or assumed) of the size of the stimulus object. Surprising as some of these phenomena seem to be, they all make sense in terms of a functional view of the nature of perception. Organisms perceive objects in space. Such a perception represents the response of the organism to a complex of stimuli, integrated with the remaining effects of past experience with similar objects. To what extent the organization of the perception depends on native or acquired organization of the nervous system is not crucial to the present discussion; suffice it to say that the continued existence of the organism is proof that its perceptions have corresponded reasonably well to the "real" world.

Let us consider the recent experiments of Hastorf (6) done at Princeton. Hastorf had some *Ss* set the size of a projected disc of light so that it represented a Ping-pong ball at the same distance as a specified marker. In a second session *S* made similar settings—"on an object that can be seen as a billiard ball." The actual settings in each case corresponded fairly well to the appropriate retinal size, or visual angle. At the end of

the second session *E* set the size of the disc at the average value obtained from *S* during the first (Ping-pong ball) session, and told him to estimate its distance relative to the marker. The disc was reported as beyond the usual marker. Another group started with "billiard balls," and two more groups set rectangles to represent calling cards or envelopes; the results from all groups were consistent.

To understand these results, it is necessary to consider the simple geometry of the situation. Perhaps it can be handled most clearly in terms of similar triangles. Let *A* represent the size of the object and *a* the size of its retinal image. The distance from *A* to the nodal point, *N*, just behind the lens of the eye, is *D*, and the distance from *N* to the image, *a*, is represented by *d*. We then have the equation:

$$A/D = a/d. \quad (1)$$

But *d*, the distance from nodal point to retina, is essentially constant, so it may be assigned a value of 1.0. The equation becomes:

$$a = A/D. \quad (2)$$

The geometry deals with "real" objects and distances; in psychological experiments, *A* and *D* become perceived size and distance, respectively.

In the first session, Hastorf's *Ss* were given *A* (Ping-pong ball size) and *D* (the distance to the specified marker), and were forced to "solve" for *a* (size of retinal image). Regardless of what they thought they were doing, the only variable they could manipulate was the size of the stim-





FIG. 1. The geometry of visual size and depth.  $A$  and  $a$  represent the size of object and retinal image, respectively.  $D$  and  $d$  represent the distance from Nodal Point ( $N$ ) to object and retina, respectively. Since  $d$  is constant, the equation may be written  $a = A/D$ . The fraction  $A/D$  is the tangent of the visual angle ( $V$ ).

ulus disc, with the resulting change in the retinal image. Similarly in the first part of the second session they were given a new  $A$  (billiard ball) and they responded with a larger setting for  $a$ . Then, without changing the instructions as to  $A$ , they were given a smaller  $a$ . The only way to keep the equation balanced was to increase  $D$ ; this is what  $S$  reported he perceived.

In the Hastorf experiment one of the factors,  $a$  (retinal size), was under direct stimulus control. A second factor,  $D$  (distance of perceived object), was controlled in a very indirect manner; it was related to the distance of another object in the visual field, which in turn was localized by the complex of stimulus and interpretative factors that enter into normal depth perception. The third factor,  $A$  (size of perceived object), was determined by a verbal reference to an object of familiar size. Perhaps the most surprising thing about the experiment is that  $S$  could make settings in such a complicated situation. The key to the explanation rests in the fact that people are so constituted that they perceive *objects localized in space*; they do not see free-floating sensations. In short, they "solve" the equation as best they can. If values for two of the factors,  $A$  and  $D$ , are given in an unequivocal manner by stimulus and/or experiential

factors,  $S$  can solve for the third factor,  $a$ , rather easily; lacking precise means of determining  $A$  and  $D$ ,  $S$  grasps at any clues available, such as instructions or sets. In other words, it is the absence of stronger clues that throws emphasis on the influence of verbal instructions.

In some of the other Ames-Cantril demonstrations the opposite situation holds true;  $S$  is given normally-reliable clues to determine all three terms of the equation, but the equation does not balance. In this case,  $S$  discards the value which seems to be least reliable. Retinal size is pretty stable; it is "stimulus-bound." Which of the other two will be discarded depends on the specific situation. For example, consider the distorted room, viewed monocularly through a peep-hole. The room is so constructed that all relations and distances in the retinal image correspond to those which would result from a normal model room, viewed from the predetermined point. In other words, a large number of parallel equations solve themselves nicely in terms of a normal rectangular room. But actually the far right-hand corner is only a few feet from the peep-hole, while the far left-hand corner is twice as far away. The only clues  $S$  has on which to judge these two  $D$ -values are accommodation and the associated sharpness of objects. These are notoriously unreliable clues (10). Hence  $S$  discards  $D$ -values based on these clues and accepts new  $D$ -values that are consistent with  $a$ - and  $A$ -values based on normally-reliable retinal distances and certain assumptions as to the rectangularity of rooms. As a matter of fact, even the marked retinal disparity produced by aniseikonia or aniseikonic lenses may not be strong enough to make  $S$  discard his assumption that room walls are rectangular;

hence the "leaf room," which is actually rectangular, but might be an irregular leafy bower, making  $A$  indeterminate, is used in testing anisokonia (1).

To the reader this discussion may sound like the resurrection of Helmholtz' "Unconscious Inference" (7). It is. This much-misunderstood concept is a very useful one in describing and predicting the perceptions of size and distance that go on endlessly in daily life; the trouble comes when one tries to use the concept to explain the mechanism of these perceptions.

Finally, it may be of interest to see how the basic equation applies to more traditional experiments. It is perfectly clear that Emmert's Law of the projected size of the visual after-image is a special case of this equation (3, 4). Here  $a$  is the retinal area subjected to light,  $D$  is the distance to which the after-image is projected, and the equation may be solved for  $A$ , the apparent size of the after-image. In one version of the size-constancy experiment (4, 8)  $a$  (retinal size) is held constant,  $D$  (distance) is varied, and the resulting changes in  $A$  (apparent object size) are recorded. Conversely,  $A$  may be kept constant, while  $a$  is adjusted to compensate for the independent variable,  $D$ .

Turning to depth perception experiments, one might expect to find many in which either  $a$  or  $A$  is kept constant, while the other of these is the independent variable, and  $D$  is the dependent variable. This seems to be too obvious an experiment for serious consideration. Most experiments in this field try to keep both  $A$  and  $a$  constant or indeterminate in an effort to measure the influence of a single clue on  $D$ . For example, a long series of experiments was devoted to the effect of varying accommodation and/or convergence on depth per-

ception (10). The present writer has already commented on the danger of keeping a determiner at a constant value (9); it requires considerable ingenuity on the part of  $E$  to set up a situation that will adequately isolate the influence of a single determiner on  $D$ .

In summary, it may be said that a large number of experiments in visual perception of size and distance, including most of the Ames demonstrations, fit neatly into the following equation:

Size of Retinal Image = Size of Object/Distance of Object. This is simple geometry. When a person is brought into an actual situation, the first of these variables, Size of Retinal Image, is essentially a sensory process; *i.e.*, it is stimulus-bound. The other two variables, Object Size and Object Distance, are often called perceptual because they are determined in a complex fashion by monocular and binocular clues, knowledge of size, etc. Where neither Size nor Distance is very well specified by clues,  $S$  will grasp at anything available to tie one of them down. If, on the other hand, both are specified, but are given values that yield a quotient inconsistent with Size of Retinal Image, one of the two perceptual variables will be discarded. Many of the Ames demonstrations show that Object Size (including shape<sup>1</sup>) is surprisingly stable in these ambiguous situations.

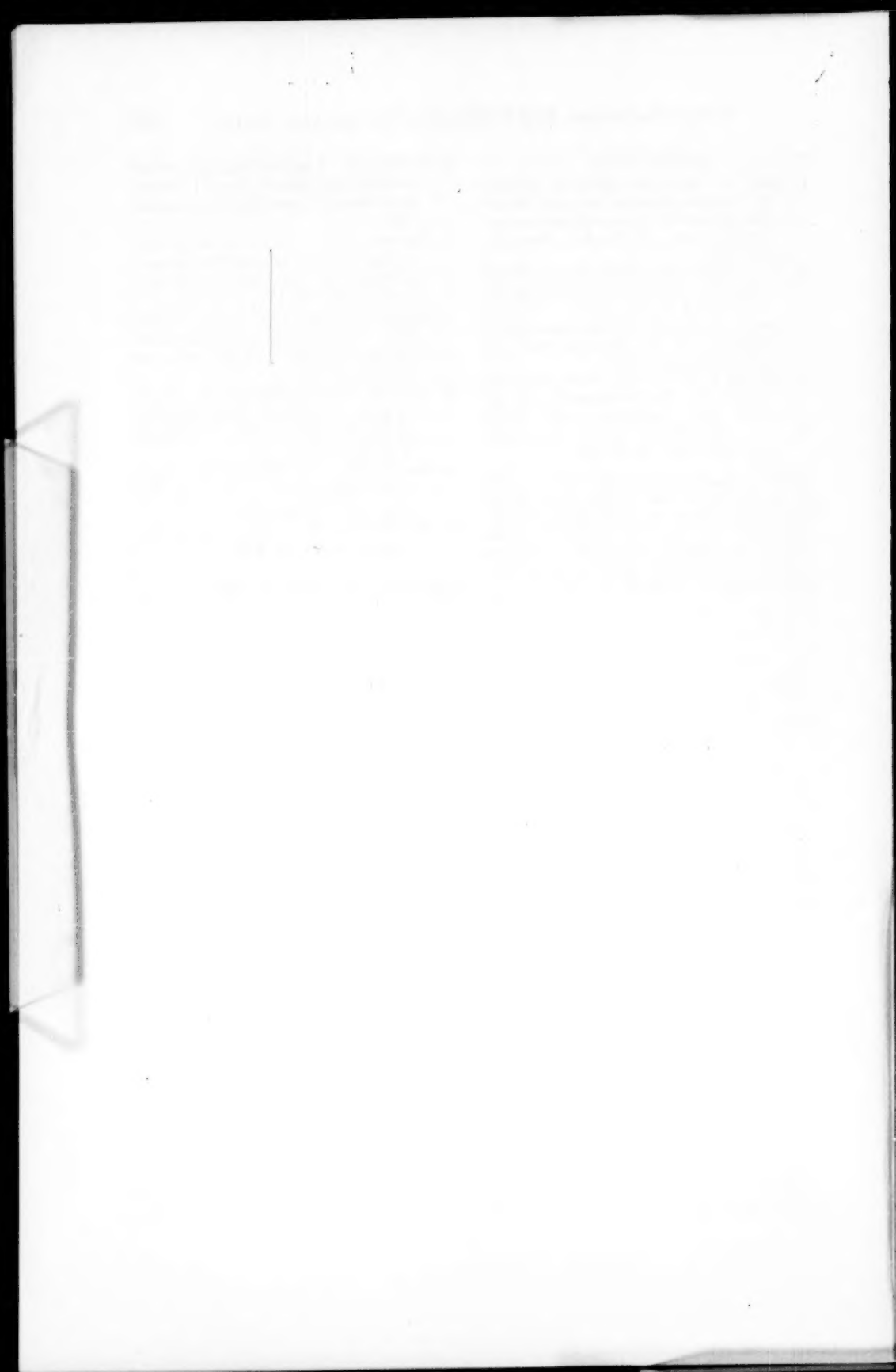
<sup>1</sup> In the interests of simplicity of exposition we have stressed only one characteristic of the object, its size. It is possible to reduce shape to the size of lines connecting pairs of salient points in the object, as the sides and diagonals of a quadrilateral. One of the most effective clues in the Ames demonstrations, interposition of objects, is a special case of shape; the retinal image of the more distant object does not correspond to the known (or assumed) shape of this object. But other related clues, as shading, cannot readily be reduced to size. If we generalize the equation by sub-

## REFERENCES

1. AMES, A. Binocular vision as affected by relations between uniocular stimulus-patterns in commonplace environments. *Amer. J. Psychol.*, 1946, **59**, 333-357.
2. ——. Some demonstrations concerned with the origin and nature of our sensations (what we experience). A laboratory manual (preliminary draft). Hanover, N. H., Hanover Institute, 1946 (mimeographed).
3. BORING, E. G. *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century, 1942.
4. ——. The perception of objects. *Amer. J. Phys.*, 1946, **14**, 99-107.
5. CANTRIL, H. *Understanding man's social behavior* (preliminary notes). Princeton: Office of Public Opinion Research, 1947.
6. HASTORF, A. H. The influence of suggestion on the relationship between stimulus size and perceived distance. *J. Psychol.*, 1950, **29**, 195-217.
7. HELMHOLTZ, H. *Physiological optics. III*. (Translated by J. P. C. Southall.) Ithaca, New York: Optical Society of America, 1925.
8. HOLWAY, A. H., & BORING, E. G. Determinants of apparent visual size with distance variant. *Amer. J. Psychol.*, 1941, **54**, 21-37.
9. SCHLOSBERG, H. Stereoscopic depth from single pictures. *Amer. J. Psychol.*, 1941, **54**, 601-605.
10. WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1938.

stituting "characteristics" for "size," it will handle all depth clues except direct physiological ones (disparity, convergence, accommodation). In cases like interposition and shading the equation will yield a relative rather than an absolute value for  $D$ , as "The Jack of Spades is behind the Queen."

[MS. received March 30, 1950]



# MEMBERSHIP RULES

## IN THE

### AMERICAN PSYCHOLOGICAL ASSOCIATION

---

There are three classes of membership in the American Psychological Association: Associate, Fellow, and Life Member.

A. The largest class of membership is *Associate*. In order to qualify as an Associate an applicant must meet one of three sets of requirements:

1. He must have a doctor's degree based in part upon a psychological dissertation and conferred by a graduate school of recognized standing; or

2. He must have completed two years of graduate work in psychology at a recognized graduate school and be devoting full time to work or graduate study that is primarily psychological in character; or

3. He must have completed one year of graduate study plus one year of professional work in psychology and be devoting full time to work or graduate study that is primarily psychological in character.

Scientists, educators, or other distinguished persons who are known to and recommended by the Board of Directors may also be elected as Associates of the American Psychological Association.

Applicants must have their applications complete by September 15. New Associates are elected in the fall and their membership will be dated as of the next year. Journals due members will begin with the January issues. Annual dues for Associates are now \$12.50.

B. Properly qualified Associate members may, upon nomination by one of the Divisions and election by the Council of Representatives, become *Fellows* of the American Psychological Association. Fellows must previously have been Associates. They must have a doctor's degree and at least five years of acceptable professional experience beyond that degree. They must be primarily engaged in the advancement of psychology as a science and a profession. Annual dues for Fellows are now \$17.50. In the American Psychological Association, no one is made a Fellow except at his own request.

C. *Life Membership* is open to members who have reached the age of 65 and who have been members for twenty years. They are exempt from dues, and receive the *American Psychologist* and the *Directory*.



# **JOURNAL OF CONSULTING PSYCHOLOGY**

*Issued bi-monthly*      *\$5.00 per year*

**American Psychological Association**

1515 Massachusetts Avenue N.W.

Washington 5, D.C.